# Ground-State Proton Transfer Kinetics in Green Fluorescent Protein

Luke M. Oltrogge, Quan Wang, and Steven G. Boxer*

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

**S** *Supporting Information*

**ABSTRACT:** Proton transfer plays an important role in the optical properties of green fluorescent protein (GFP). While much is known about excited-state proton transfer reactions (ESPT) in GFP occurring on ultrafast time scales, comparatively little is understood about the factors governing the rates and pathways of ground-state proton transfer. We have utilized a specific isotopic labeling strategy in combination with one-dimensional $^{13}C$ nuclear magnetic resonance (NMR) spectroscopy to install and monitor a $^{13}C$ directly adjacent to the GFP chromophore ionization site. The chemical shift of this probe is highly sensitive to the protonation state of the chromophore, and the resulting spectra reflect the thermodynamics and kinetics of the proton transfer in the NMR line shapes. This information is complemented by time-resolved NMR, fluorescence correlation spectroscopy, and steady-state absorbance and fluorescence measurements to provide a picture of chromophore ionization reactions spanning a wide time domain. Our findings indicate that proton transfer in GFP is described well by a two-site model in which the chromophore is energetically coupled to a secondary site, likely the terminal proton acceptor of ESPT, Glu222. Additionally, experiments on a selection of GFP circular permutants suggest an important role played by the structural dynamics of the seventh $\beta$-strand in gating proton transfer from bulk solution to the buried chromophore.
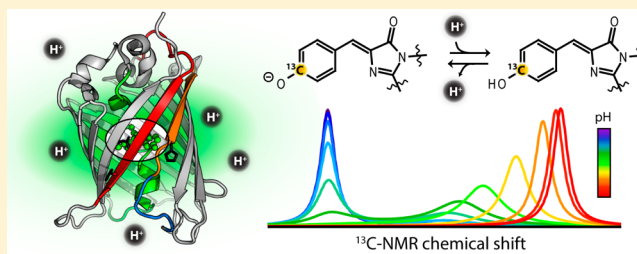
The green fluorescent protein (GFP) and its variants have found broad application in molecular and cellular biology for fluorescence imaging and chemical sensing. This widespread use is owed in large part to the richness and tunability of the photophysics and photochemistry of the autocatalytically formed chromophore.[1] Proton transfer is an important player in this respect and is central to the use of fluorescent proteins (FPs) as pH indicators[2,3] and as reversibly photoswitchable dyes utilized in super-resolution microscopy.[4] A better understanding of the factors influencing the rates and pathways of proton transfer in FPs in both the ground and excited states can aid in rationally developing increasingly robust probes and, more generally, offer insight into proton transfer in biology. In the following, we introduce a $^{13}C$ NMR probe into the chromophore whose chemical shift is highly sensitive to the protonation state of the chromophore and to the dynamics of ground-state proton transfer.

The GFP chromophore is comprised of a 4-hydroxybenzylidene imidazolinone structure formed from a self-catalyzed cyclization, dehydration, and oxidation of residues Ser65, Tyr66, and Gly67. The chromophore is located within the central $\alpha$-helix and surrounded by 11 $\beta$-strands that form the compact $\beta$-barrel fold (Figures 1 and 2A). Many mutations that affect the absorbance and fluorescence spectra, the quantum yield, and the isomerization efficiency via direct and indirect interactions with the chromophore have been identified. The most dramatic examples of these effects involve those mutations that perturb the chromophore ionization state.

The GFP chromophore is known to exist in a protonation–deprotonation equilibrium at the phenoxy position, which manifests itself as two electronic absorbance bands in the visible: a neutral form (A state) at ~395 nm and an anionic form (B state) at ~470 nm. Interestingly, the ratio of the A and B states in wild-type GFP (wtGFP) is largely independent of solution pH over a wide range (~6–12). Upon excitation of the A state, wtGFP exhibits fluorescence emission with an anomalously large apparent Stokes shift. Early work in our lab established the origin of this behavior in an excited-state proton transfer (ESPT),[5] and subsequent structural and spectroscopic experiments by others revealed the internal proton transfer network through which this proton is shuttled.[6,7] In contrast to wtGFP, most of the commonly used variants (EGFP, YFP, Superfolder, etc.) interrupt this ESPT pathway and the A-state excitations result in neither ESPT nor direct A* emission. Moreover, in these variants, the ratio of the A to B states can be titrated by the solution pH.[8] The titration curves resulting from many FP variants, however, do not exhibit simple single-site behavior and can show features such as negative cooperativity and spectra with mixed states in the limits of high or low pH. These nonideal situations have been attributed to electrostatic coupling between the chromophore and additional ionizable sites in its proximity. A two-site thermodynamic model has been shown to give satisfactory agreement across the range of titratable FPs and can even account for the pH insensitivity of the wild-type protein.[9,10]
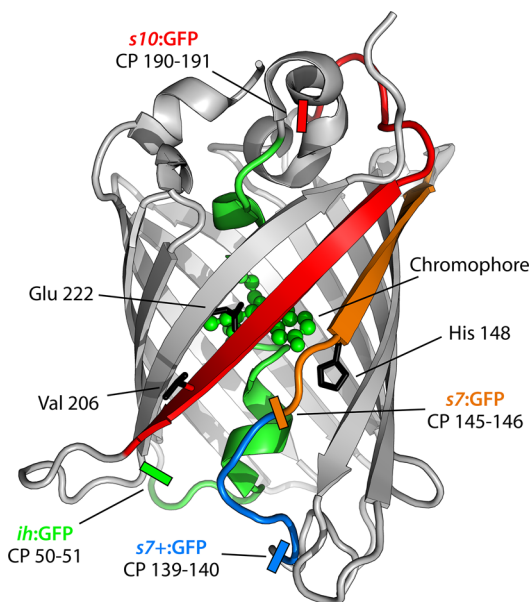
**Figure 1.** Ribbon structure of Superfolder GFP (Protein Data Bank entry 2B3P[30]). The central helix and β-strands 7 and 10 are colored green, orange, and red, respectively. Important residues are labeled and displayed as black sticks. Circular permutation sites are indicated by colored rectangles and the residue numbers between which the new termini were located. For all circular permutants, the native N- and C-termini are linked by the sequence GGTGGS. Image created with PyMOL (Schrödinger, LLC).
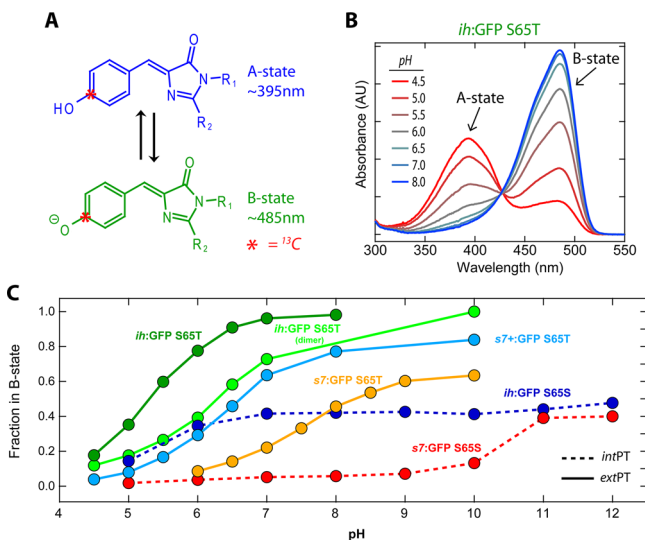


**Figure 2.** (A) Protonated and deprotonated states of the GFP chromophore result in two visible absorbance bands that can be titrated by pH. The carbon atom that originates from $[^{13}C_\zeta]$-L-Tyr66 and has the $^{13}C$ label is highlighted with a red asterisk. (B) UV–vis absorbance pH titration of *ih*:GFP S65T showing the electronic absorbance bands for the A and B states. These spectra are representative of all variants presented. (C) Fractions of the population in the B state vs pH for a selection of mutants and circular permutants. The titrations demonstrate behavior characteristic of coupled sites with apparent negative cooperativity and some values not converging to 100% B state in the high-pH limit. *ih*:GFP S65S and *s7*:GFP S65S are classified as *int*PT, while all others are classified as *ext*PT.

The kinetics of proton transfer in FPs has been a subject of extensive study. The PT is important for many FP applications,

and the unique chromophore provides a convenient optical handle by which to follow PT within proteins in a more generic sense. Ultrafast pump–probe experiments have elucidated the details of ESPT in wtGFP[6] and have provided information about the mechanism of reversible photoswitching in photo-activatable FPs such as Dronpa.[11,12] Ground-state proton transfer has also been investigated using techniques of fluorescence correlation spectroscopy (FCS),[13−16] pump–dump–probe transient absorption,[17] nanosecond pH-jump kinetics,[18−20] and stopped-flow fluorescence.[3,21] Many of these data have been interpreted through the single-site titration model and have led to a relatively wide range of estimates for the protonation and deprotonation rate constants. For example, the microscopic deprotonation constant for the widely used EGFP variant has been variously reported to have values as low as 45 s$^{-1}$[19] and as high as 4500 s$^{-1}$.[13]

The chromophore is cradled within a compact β-barrel in which it is physically isolated from the solution as evidenced by the inefficacy of external fluorescence quenchers.[22] A recurring theme then, in previous publications, has been to rationalize the communication of the solution pH to the chromophore via gross protein dynamics, specific residue-mediated proton transfers, or some combination of the two.[19,20] Strand 7 and, in particular, residue His148 have received attention as a likely "gatekeeper" of PT.[20,23] Crystal structures show that His148 is directly hydrogen-bonded to the ionizable phenoxy of the chromophore;[8] consistent with this, NMR studies have revealed strand 7 to be particularly dynamic from H/D exchange experiments relative to the other β-strands.[24]

NMR has long been used as a powerful tool to assess residue-specific pH titrations within complex biomolecules.[25−27] In this study, we have focused on direct $^{13}C$ NMR spectroscopy of the GFP chromophore as a method possessing high resolution, chemical specificity, sensitivity to the chemical exchange time scale, and independence from, but potential correlation with, the optical properties. Recent work in our lab and others has utilized direct carbon detection of tyrosine isotopically labeled at position ζ [adjacent to the hydroxyl group (see the asterisks in Figure 2A)] to study the thermodynamic properties and qualitative kinetics of PT.[26,28,29] After the GFP chromophore maturation process, $^{13}C_\zeta$ of Tyr66 is located directly adjacent to the ionizable site and provides a convenient and non-perturbative spectator with which to follow PT. We utilize this approach to look at the widely used Superfolder GFP,[30] a particularly stable and fast-folding GFP variant, and some select circular permutants (Figure 1) to analyze the interplay between structural rigidity and PT from this new nonoptical observable. Moreover, we combine the information from NMR with absorbance and fluorescence pH titrations and FCS measurements to offer a more complete picture of the underlying mechanism.

## ■ MATERIALS AND METHODS

**Protein Expression and Purification.** All circular permutant genes were designed by us and then synthetically prepared by Genscript. The genes were inserted into the pET-15b vector along with hexahistidine tags and were cloned into BL21(DE3) *Escherichia coli* (Invitrogen). The point mutations and insertion mutations were performed with the Stratagene QuikChange mutagenesis kit according to the manufacturer's protocol.

The proteins were expressed from cells grown in minimal M9 salts medium supplemented with each of the common L-

amino acids (Sigma) at 25 mg/L with the exception of tyrosine, which was replaced with $[^{13}C_\zeta]$-L-tyrosine (Cambridge Isotopes, 95−99% $^{13}C$). The cells were grown at 37 °C until the $OD_{600}$ reached 0.4−0.6 before being induced with 0.25 g/L isopropyl β-D-1-thiogalactopyranoside. The cells were harvested 4 h after induction and lysed. The supernatant was purified by nickel affinity with Ni-NTA agarose resin (Qiagen) and further with anion exchange chromatography (HiTrap Q HP column, GE Healthcare). The purity of all the proteins was verified via electrospray ionization mass spectrometry.

**Absorbance and Fluorescence pH Titrations.** The pH titrations were performed using a master buffer consisting of 20 mM phosphate, 30 mM citrate, and 150 mM NaCl, each adjusted to the desired pH with additions of a HCl or NaOH solution. The proteins were exchanged with these buffers and the absorbance spectra collected at approximately 2 $\mu$M with a PerkinElmer Lambda 25 UV−vis spectrometer. The fluorescence spectra were collected at protein concentrations of 0.1− 0.2 $\mu$M in right angle detection mode on a PerkinElmer LS 55 fluorescence spectrometer.

**One-Dimensional (1D) $^{13}C$ NMR Measurements.** The protein was concentrated to ∼2 mM using 10 kDa molecular mass cutoff spin filters (Amicon), made to 7% $D_2O$ for locking, and loaded into a 5 mm Shigemi low-volume NMR tube. 1D $^{13}C$ NMR spectra were all recorded with $^1H$ decoupling on a Varian Inova 300 MHz spectrometer unless otherwise noted. The $T_1$ relaxation time was measured for the chromophore $^{13}C_\zeta$ resonance using standard inversion recovery techniques. The best duty cycle in the 1D spectra was achieved empirically using a 50° pulse and a delay time of 4.0 s. All experiments were performed with the temperature control set to 25 °C. The typical samples for NMR contained 40 mM buffer (HEPES for high-pH samples and MES for lower-pH samples) and 40 mM NaCl. Phosphate and citrate buffers and higher salt concentrations were also used for a limited set of spectra and found to give identical results for a given pH.

**Saturation Transfer.** The protein *ih*:GFP S65T at pH 6.70 was irradiated on the A-state basis resonance of 162 ppm (Figure 3C) with a soft pulse for 400 ms. The free induction decay was collected following a subsequent $\pi$/2 hard pulse. The same procedure was repeated off-resonance at the same frequency difference between the A- and B-state resonances but to the opposite direction downfield (see section S.6 of the Supporting Information for additional information).

**Chemical Exchange Rate from NMR Inversion Transfer.** For the slow chemical exchange proteins, we employed the selective inversion transfer method of Robinson et al.[31] with the pulse sequence $[(\pi/2)_x(^{13}C)-\tau_1-(\pi/2)_x(^{13}C)-\tau_{mix}-(\pi/2)_x(^{13}C)]$, where $\tau_1$ is the time required for the exchanging states to precess to antiphase (440 $\mu$s in these experiments) and $\tau_{mix}$ is the variable time in which the chemical exchange is allowed to proceed (0−40 ms). This experiment was repeated in triplicate for *ih*:GFP S65T at pH 6.70.

**Fluorescence Correlation Spectroscopy.** Protein samples were exchanged into buffers with 20 mM citric acid, 20 mM sodium phosphate, 10 mM glycine, and 150 mM NaCl all adjusted to the appropriate pH by additions of 1 M NaOH. The FCS measurements were taken with protein concentrations in the range of 0.2−5.0 nM and all performed at room temperature. The sample was excited with a 488 nm argon ion laser (Innova 70, Coherent) focused to a near-diffraction-limited spot. The fluorescence was filtered from the excitation beam with a dichroic mirror (z488rdc, Chroma) together with
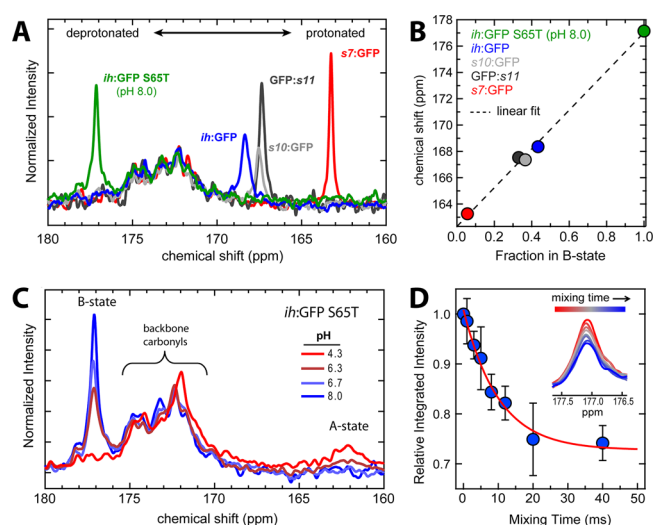


**Figure 3.** (A) 1D $^{13}C$ NMR spectra of *int*PT Superfolder circular permutants with different B-state fractions at pH 8.0. *ih*:GFP S65T at pH 8.0 is also included as a pure B-state reference. All of these proteins result in a single NMR peak for the chromophore $^{13}C_\zeta$ peak, albeit at very different chemical shifts [the other eight labeled Tyr residues clustered around 155 ppm are used for normalization (see section S.4 and Figure S1 of the Supporting Information)]. (B) Chemical shifts of the peaks plotted vs the fraction in the B state deduced from the absorbance spectrum (see Figure 2B and section S.2 of the Supporting Information). The near linearity of the points is indicative of fast exchange between protonated and deprotonated forms. Extrapolation of the fit gives 177.2 ppm as the B-state basis and 162.0 ppm as the A-state basis chemical shift. (C) 1D $^{13}C$ NMR pH titration of *ih*:GFP S65T. The peak at 177.2 ppm decreases in intensity with a lower pH accompanied by the emergence of a broad peak near 162 ppm. (D) Integrated intensities of the B-state magnetization from the selective inversion experiment with *ih*:GFP S65T at pH 6.70 as a function of mixing time with error bars representing one standard deviation from triplicate measurements. The apparent chemical exchange rate constant is 99 $s^{-1}$ at pH 6.70, which leads to a predicted deprotonation rate constant of 74 $s^{-1}$. The inset is the B-state peak at increasing mixing times.

two emission filters (HQ535/70, Chroma) and passed through a 75 $\mu$m confocal pinhole and detected with an avalanche photodiode (SPCM-AQR-14 PerkinElmer). The typical irradiance was 10 $kW/cm^2$; however, some trajectories were also recorded with an irradiance of 2.5 $kW/cm^2$ to determine the effect of the light intensity. The dimensions of the fluorescence detection volume were experimentally measured by scanning a piezo stage over a sub-diffraction-limited fluorescent bead (100 nm in diameter) and confirmed by measurements on small dyes (Atto 488) with known diffusion coefficients ($D = 400$ $\mu m^2$/ s).[32] This detection volume was determined to conform closely to a Gaussian profile in all three dimensions with a typical size of $r_0 = 0.4$ $\mu$m and $z_0 = 1.6$ $\mu$m. ACFs were calculated by an efficient photon-by-photon algorithm.[33]

■ **RESULTS**

The proteins we investigated can be broadly grouped into two categories according to the residue at position 65. In wtGFP, Ser65 has an important role in stabilizing a hydrogen bond network through which ESPT occurs. When the protein is excited at the A-state band (∼395 nm) or B-state band (∼470 nm), both give rise to fluorescence that peaked at ∼505 nm. Furthermore, the ratio between the A and B states, based on

their characteristic absorption bands, is relatively constant over a wide pH range. Superfolder GFP with Ser65 also displays these two features, including ESPT occurring in a matter of picoseconds as measured by ultrafast fluorescence upconversion (data not shown). In contrast, the original Superfolder GFP has Thr65.[30] This mutation causes the chromophore to exist primarily in the deprotonated form at neutral pH and leads to seemingly single-site titration behavior with an apparent $pK_a$ of ~5.7 (see Figure 2B,C). The A state that grows in at low pH is nonfluorescent because of efficient nonradiative decay of A*, and it is incapable of ESPT because of the disruption of the proton transfer wire by Thr65. We focus on these two classes of proteins because the Ser65 group exemplifies internal proton transfer (*int*PT) along a well-characterized proton transfer pathway, while the Thr65 group is in clear chemical communication with the solution via external proton transfer (*ext*PT).

All circular permutants of Superfolder GFP bearing the Ser65 reversion presented herein show the characteristics diagnostic of internal proton transfer. More specifically, all have A:B ratios that are largely independent of pH, and both steady-state and time-resolved fluorescence spectroscopy reveal ESPT with the same kinetics and isotope effect that were observed in wtGFP, strong evidence of functional and structural similarity. Interestingly, the circular permutants that we have made differ in their A:B ratios. This unanticipated fact, discovered in the course of studies of split GFPs,[34,35] has provided a convenient way to span the range of protonation states while maintaining internal PT. Throughout the text, the naming convention for the circular permutants is to give the N-terminal structural element (*s1−s11* for the strands and *ih* for the interior helix) followed by a colon and GFP; e.g., *s7*:GFP is the protein with native termini fused and the seventh β-strand at the new N-terminus (see section S.1 of the Supporting Information for all protein sequences). In the construct *s7+*:GFP, the plus sign is used to distinguish a circular permutation site that includes a larger portion of strand 7. The locations of all circular permutants used are indicated in Figure 1.

The absorbance spectra as a function of the pH in the S65T *ext*PT circular permutants were consistent with previous measurements of S65T GFPs in that they displayed A:B peak ratio changes and clean isosbestic points,[3,8] e.g., *ih*:GFP S65T in Figure 2B, a clear indication of exactly two optical states. The fraction of the population in the B state at all pH values was calculated using extinction coefficients for the A and B basis states derived from absorbance and fluorescence titrations (as described in section S.2 of the Supporting Information).

The GFP constructs used in these experiments all have nine tyrosines, including Tyr66, which goes on to form the hydroxybenzyl portion of the chromophore (Figure 2A). The 1D $^{13}$C NMR spectra of the all of the proteins showed a cluster of eight peaks near 155 ppm corresponding to protonated tyrosine (Figure S1 of the Supporting Information). As seen in Figure 3A, the chemical shift of $^{13}$C$_\zeta$ of the chromophore is shifted downfield away from the tyrosine cluster as a consequence of the chemical transformation involved in maturation that extends the conjugation. The very different chemical shifts for different circular permutants reflect the variations in chromophore ionization equilibria. Previous work on a model chromophore found the protonated $^{13}$C$_\zeta$ resonance at 160 ppm,[36] while a recent NMR study of photoswitching in the fluorescent protein Dronpa revealed a deprotonated chromophore $^{13}$C$_\zeta$ chemical shift of 177 ppm in the *cis*

isomer.[37] These prior assignments provide additional confidence that we are indeed observing the chromophore resonance. The somewhat structured broad peak centered at ~172 ppm is due to the ~240 backbone carbonyl carbon resonances at natural abundance and is also present in unlabeled protein samples (Figure S3 of the Supporting Information). The location and width of the unique chromophore peak depend sensitively on the A:B ratio as deduced from the absorption spectra (Figure 2B) as well as the dynamical properties of the proton transfer discussed below. These characteristics varied markedly between proteins with *int*PT versus *ext*PT and among different circular permutants.

**Internal Proton Transfer.** All of the *int*GFPs measured gave a single peak from the chromophore $^{13}$C$_\zeta$ (Figure 3A). However, the location of this peak changed from protein to protein as a function of the A- and B-state occupancies (Figure 3A,B). Also included in Figure 3A is the titratable *ih*:GFP S65T in the limit of high pH to provide the fully deprotonated basis at a chemical shift of 177.1 ppm. *s7*:GFP S65S is ~95% protonated and appears at 163.6 ppm, and linear extrapolation predicts a protonated basis of 162.0 ppm. This represents a chemical shift dispersion of >15 ppm, highlighting the sensitivity of this probe to the chromophore protonation state. Figure 3B shows that there is good linear agreement between the fraction in the B state calculated from UV−vis absorbance measurements and the $^{13}$C NMR chemical shifts. This single-peak behavior is indicative of fast exchange on the NMR time scale in which only one peak is observed at a population-weighted chemical shift value. For reference, the data in Figure 3A were collected with a 300 MHz NMR spectrometer, which makes the inverse of the frequency difference between the basis states approximately 1 ms. Systems undergoing much faster chemical exchange will appear as a single peak, while those exchanging much slower would result in two peaks.

**External Proton Transfer.** The protein *ih*:GFP S65T was selected as the model for external proton transfer. It exhibits fluorescence properties identical to those of the unpermuted protein and has simple single-site titration behavior with an apparent $pK_a$ of 5.9 (see Figure 2B,C, dark green). 1D $^{13}$C NMR spectra were collected over a range of pH values. In contrast to the *int*PT samples, in which the chemical shift was proportional to the ionization equilibrium, these spectra showed a single peak at the deprotonated basis frequency that did not shift but rather underwent a decrease in intensity with lower pH (Figure 3C). At the lowest pH values, another much broader peak begins to appear around 162 ppm, at the location predicted from the linear extrapolation for the protonated form of the chromophore from Figure 3B. This behavior suggests that the system is undergoing slow chemical exchange on the NMR time scale.

To verify in the *ih*:GFP S65T sample that the broad putative A-state peak at 162 ppm was in fact in chemical exchange with the peak at 177 ppm (Figure 3C), we performed saturation transfer experiments in which a soft 400 ms saturating pulse was applied at 162 ppm and another with the pulse at an equivalent distance upfield at 195 ppm. If, over the duration of the saturation pulse, chemical exchange were occurring between states with chemical shift bases of 177 and 162 ppm, then the intensity of the peak at 177 ppm should decrease because the population transiently converts its chemical shift to the frequency of irradiation. Indeed, only the on-resonance pulse caused the B-state peak to disappear, while the off-resonance

pulse had no effect (Figure S6 of the Supporting Information). The slight changes occurring within the carbonyl band with pH are likely due to local perturbations due to additional ionizable groups throughout the protein (for additional details, see section S.6 of the Supporting Information).

To push beyond the more qualitative line shape and population connectivity analysis, we turned to time-resolved NMR. The kinetics of proton transfer were measured using the selective inversion transfer experiment introduced by Robinson et al.[31] In particular, we investigated the *ih*:GFP S65T protein as it was our best model system for slow exchange. In brief, this technique prepares the spin system in such a way that the magnetization vectors caused by the respective populations of the A and B states are antiphase and parallel to the permanent magnetic field. During a subsequent mixing time prior to the acquisition pulse, the populations undergo chemical exchange and interconvert, thereby lowering the magnitude of the net magnetization vectors due to each state. This effect is manifested as a decrease in the magnitude of the well-resolved B-state peak as the mixing time delay grows larger (Figure 3D).[a]

The experiments were performed in triplicate at pH 6.70 where an appreciable A-state population exists and at pH 8.00 as a control for ∼100% B state (see Figure 3D and Figure S7 of the Supporting Information). The chemical exchange mixing times were varied from 0 to 40 ms. The pH 6.70 data revealed a decay constant of 99 s$^{-1}$ (a lifetime of 10.1 ms). Bootstrap analysis was performed in which sets of data were randomly selected from the triplicate measurements and gave a range of 58−113 s$^{-1}$ in the 90% confidence interval.[38] In combination with the fraction deprotonated obtained from UV−vis spectroscopy, this yielded a predicted unimolecular deprotonation constant of 74 s$^{-1}$. The same experiment at pH 8.00 for chemical exchange mixing times of 0 and 40 ms revealed no decrease in intensity. This result is consistent with expectations because the titration shows that the fraction of deprotonated chromophore is >99%. Additional data were taken at pH 6.30 and show a faster decay rate and with a larger net change in the B-state peak size with increasing time. A global model using a deprotonation rate constant of 74 s$^{-1}$ and deprotonated fractions ($p_B$) from UV−vis absorbance gave close quantitative agreement in the relative B-state magnetization magnitudes at all chemical exchange mixing times (Figure S7 of the Supporting Information) (for details on equations and fitting, see section S.7 of the Supporting Information).

As shown in Figure 4, another *ext*PT protein, *s7*:GFP S65T, had a distinctly different pH response. For this construct, the chromophore contributed only a single peak. Furthermore, the chemical shift of the peak was sensitive to the pH and shifts in accordance with the titration population changes (Figure 4A). As with the *int*PT proteins, this behavior indicates that the system is also undergoing fast exchange on the NMR time scale. Interestingly, when the strand 7 circular permutation site is relocated six residues closer to the N-terminus, a construct we name *s7+*:GFP S65T (see Figure 1), the NMR spectra are remarkably different. Unlike the nonextended *s7*:GFP S65T that shows a single peak, this species has two clear peaks, one at the deprotonated basis of 177 ppm that undergoes a decrease in intensity as the pH is lowered and another further upfield whose chemical shift is sensitive to pH (see Figure 4B). The fact that the chromophore spectrum has features simultaneously exhibiting characteristics of both fast and slow exchange (cf. panels A and C of Figure 3, respectively)
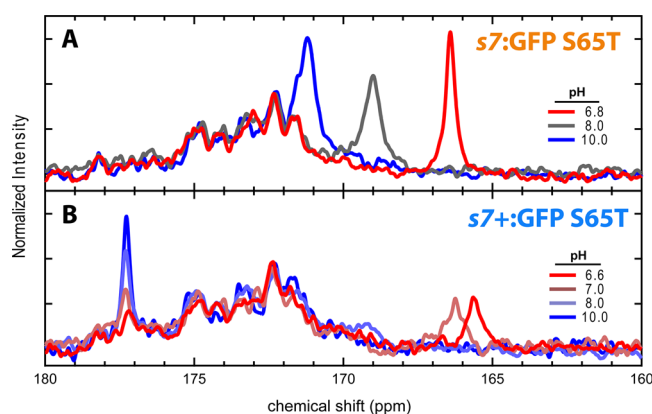


**Figure 4.** (A) 1D $^{13}$C NMR pH titration of *s7*:GFP S65T. The single peak with a chemical shift that changes as the pH increases indicates fast exchange. (B) 1D $^{13}$C NMR pH titration of *s7+*:GFP S65T. The peak at 177 ppm with a changing intensity is characteristic of slow exchange; however, another peak first appearing as a broad feature near 169 ppm at pH 8.0 and moving toward ∼165 ppm at pH 6.6 suggests a multisite model with fast exchanging components.

strongly suggests that a simple single-site model is inadequate to account for the proton transfer dynamics. The same saturation transfer procedure described above was employed for *s7+*:GFP S65T at pH 7.0 and likewise revealed population connectivity between the two peaks (see Figure S6 of the Supporting Information). This provides clear evidence that we are indeed observing interconverting protonation states.

**Fluorescence Correlation Spectroscopy (FCS) Measurements.** FCS monitors stochastic fluorescence trajectories from a small illuminated volume and via autocorrelation analysis reveals information about the diffusional and internal dynamics of an ensemble of fluorophores. Like NMR, it is an equilibrium fluctuation technique. We performed FCS on all of the proteins, and as prior investigators have done, we attribute the decays in the autocorrelation function to three primary sources: (1) diffusion out of the focal volume, (2) light intensity-dependent dark states (e.g., triplet states), and (3) pH-dependent protonation−deprotonation chemical exchange. Autocorrelation functions (ACFs) were obtained for each protein over a range of pH values to determine the apparent chemical exchange rates (see section S.3 of the Supporting Information).

The results for all *ext*PT proteins were very much in line with those of similar experiments performed on GFP S65T, EGFP, Superfolder GFP, and others.[13−15] The ACF decay became markedly faster at lower pH, indicating faster chemical exchange for higher proton concentrations (Figure 5). Specifically, *ih*:GFP S65T (Figure 5A) had an apparent deprotonation rate constant of 3100 s$^{-1}$ (see section S.3 of the Supporting Information). This is a 42-fold increase over the rate measured by dynamic NMR (74 s$^{-1}$) that should, in principle, be reporting on the same chemical states.

The wide discrepancy between these results induced us to further explore potential artifacts from which either technique may suffer. The two most significant differences between the experimental conditions are the concentration (∼500 pM for FCS and 1−2 mM for NMR) and the visible light illumination (∼10 kW/cm$^2$ for FCS and 0 W/cm$^2$ for NMR). The concentration-dependent oligomerization of FPs has long been appreciated and raised the possibility that the slow dynamics observed via NMR were a result of occlusion of the proton
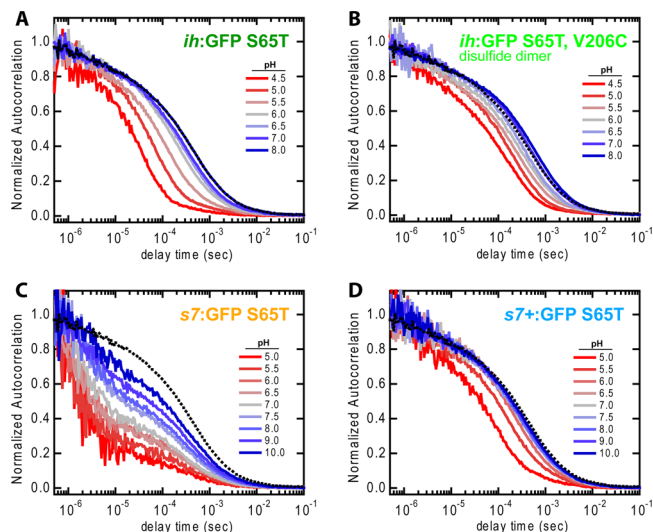
**Figure 5.** FCS ACFs for a selection of *ext*PT proteins and a range of pH values. The dashed black line is *ih*:GFP S65T at pH 8.0 and is included as a reference for the typical monomer ACF in the absence of protonation—deprotonation flickering. The dimer in panel B is included for a more direct comparison with the NMR results for *ih*:GFP S65T (see section S.5.a of the Supporting Information). (C) *s7*:GFP S65T is notable for its heterogeneous ACF decay with components occurring significantly faster than the other species.

entry point due to protein dimerization. Furthermore, under the intense illumination in the FCS focal volume, there could be some light-driven dynamics. We performed a series of control experiments to test these possibilities. Briefly, the concentration issue was addressed both by creating a covalent dimer for FCS measurements and by performing NMR at low concentrations with a dimerization-suppressing mutant. The light effect was likewise approached from both angles by changing the illumination intensity in the FCS experiments as well as measuring NMR spectra with *in situ* laser illumination via an optical fiber inserted into the bore of the magnetic in a manner similar to that described by Mizuno et al.[37] In summary, it was found that neither of these factors can account for the large rate differences (see section S.5 of the Supporting Information for details).

## ■ DISCUSSION

The photophysical behavior of fluorescent proteins is quite diverse and exhibits a hierarchy of time scales (from picoseconds to minutes) and sensitivity to a range of external factors, including light intensity, ionic strength, and pH. Whereas a majority of studies concerning chromophore dynamics have used optical observables, we have complemented the optical techniques of absorption, fluorescence, and FCS with site-specific NMR spectroscopy to probe ground-state proton transfer. The results that we obtained indicate that the dynamics of dark PT are inconsistent with a single-site titration and imply kinetic and thermodynamic coupling between multiple sites. Protein variants differing by mutation or circular permutation were shown to diverge widely in their dynamic characteristics, many of which appeared to disagree internally. However, as described below within a two-site titration model, these divergent features can be reproduced in a comparatively parsimonious way.

Before we undertake the construction of a semiquantitative two-site model, it is important to highlight again precisely what

needs to be modeled and which discrepancies need to be resolved. First, from a thermodynamic perspective, the model must account for the abnormal pH titration behaviors of negative cooperativity and apparent mixed-state asymptotes (Figure 2C). Second, the model should be able to unify the apparent slow proton exchange observed by NMR in certain proteins with the much faster chemical exchange indicated by FCS (compare, for example, Figure 3C,D with Figure 5A). Lastly, a successful model should be able to account for the presence of multiple chromophore NMR peaks in a single variant displaying contradictory dynamic signatures, that is, peaks simultaneously showing slow and fast exchange behaviors upon pH titration (see Figure 4B). An important assumption implicit in this treatment is that the same general kinetic form underlies all of the variants studied here. We take the sequence and functional similarities of the protein variants as justification for the use of this assumption as well as the resulting model's concise form and relative success.

**Kinetic Model of Proton Transfer.** Because neither concentration nor illumination differences were able to reconcile the large discrepancy between the NMR and FCS for the proton transfer reaction, we must consider alternative models that accommodate both sets of observables. As mentioned above, a two-site model has been successfully applied to explain the noncanonical UV−vis pH titration results and to address the kinetic results of certain pH-jump experiments.[10] We utilize such a model to critically analyze the varied phenomenology of the protein variants presented here and assess whether this framework can encompass all our observations (this two-site model is summarized in Figure 6).

The essence of the two-site model is that the behavior of the chromophore is connected to a second group, labeled "X" in Figure 6, whose ionization state influences the dynamical properties of proton transfer. The thermodynamics of this system can be conveniently and completely represented by the apparent $pK_a$ of the chromophore, the apparent $pK_a$ of the
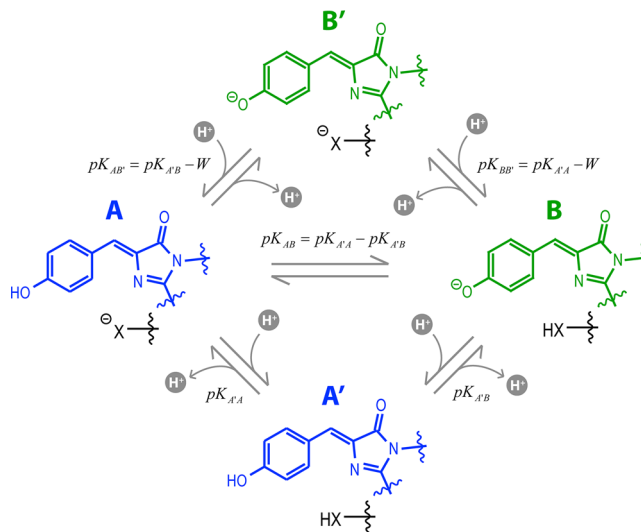


**Figure 6.** Two-site model for proton transfer. The protonated chromophore forms are colored blue and the deprotonated forms green. The second site is shown as a black "X". The microscopic site $pK_a$ values are given for each transition and include the site coupling energy ($W$) where necessary. For both NMR and FCS, only chemical exchange linked to the chromophore ionization via transitions between A states (blue) and B states (green) can be observed.
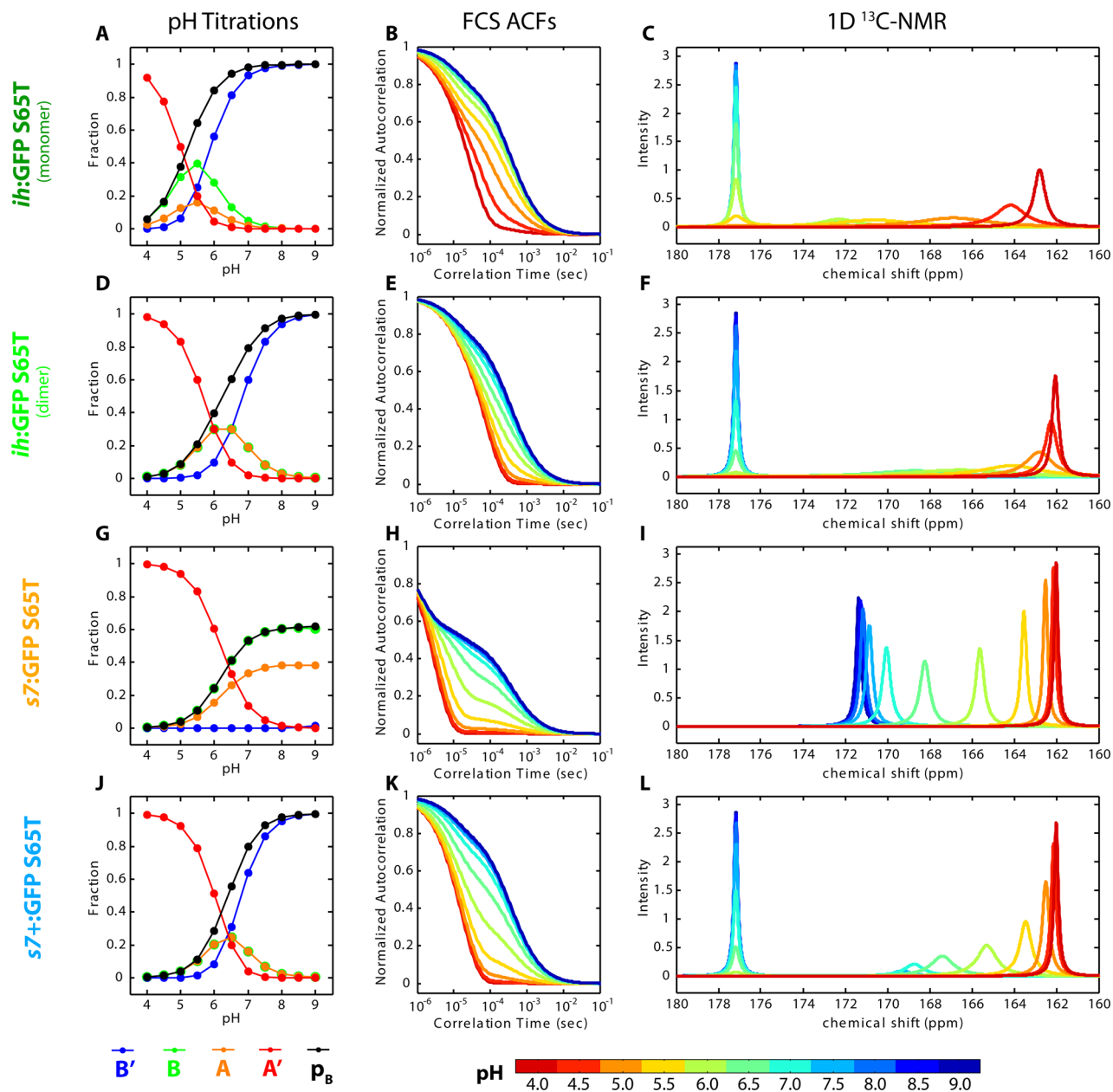
**Figure 7.** Simulated pH titrations, FCS ACFs, and 1D $^{13}$C NMR spectra for all measured *ext*PT proteins. In some cases, the B species (green) is partially obscured in the pH titrations. The NMR spectra are all normalized such that the integrated peak area is always unity.

secondary site, and a coupling energy $W$ between these two sites. Specifically, this means that the p$K_a$ of the chromophore (p$K_{A'B}$) is related to its p$K_a$ in the presence of the ionized second site (p$K_{AB'}$) as p$K_{AB'}$ = p$K_{A'B}$ − $W$. The meaning of $W$ can be rationalized as arising, in part, from the electrostatic energetic penalty of forming an anion in the proximity of another. Like the p$K_a$, $W$ is dimensionless but proportional to the standard free energy by the factor 2.303$RT$, where $R$ is the ideal gas constant and $T$ is the absolute temperature.

The UV−vis pH titration is completely specified by the thermodynamic parameters because it reports on only equilibrium concentrations. The FCS ACFs and NMR spectra, in contrast, are influenced by the system dynamics and require chemical exchange rates for prediction. Our convention is to define the kinetics of each reaction by the chemical exchange rate at a pH equal to its p$K_a$. We chose to utilize the kinetic master equation formalism and treat all reactions as

pseudounimolecular. This assumption is well justified for the experiments described herein because they are all equilibrium fluctuation measurements and therefore subject to constant pH. The calculations of predicted equilibrium pH titrations, FCS ACFs, and 1D $^{13}$C NMR spectra were performed numerically in Matlab (for additional details, see section S.8 of the Supporting Information).

**Manual Modeling.** The scheme in Figure 6 formally has ten parameters, of which eight are independent. Some of these parameters may be surmised, or at least bounded, by a careful consideration of experimental data. We found it to be convenient and intuitive to separate the parameters into two sets: those defining only the thermodynamics and those defining only the kinetics. The equilibrium UV−vis pH titration data are determined solely by the thermodynamic terms. Specifically, the thermodynamic parameters are the microscopic p$K_a$ values of the chromophore (p$K_{A'B}$) and site

"X" ($pK_{A'A}$), provided that the other site is in its neutral form, and the energetic coupling between the sites upon ionization ($W$). These three parameters together fully define the equilibrium state and thus completely specify the pH titrations. Within this convention, the kinetics are independently described for each microscopic step by the apparent rate constant at a pH equal to the $pK_a$ where the forward and reverse pseudounimolecular rate constants are equal. The single exception is the apparent internal proton transfer rate ($k_{AB}$) that is pH-independent. Collectively, all of these parameters allow the calculation of the microscopic protonation and deprotonation rate constants necessary for constructing the kinetic master equation (see section S.8 of the Supporting Information for details).

The FCS measurements are sensitive to chromophore protonation−deprotonation transitions occurring on time scales shorter than those that can be resolved by NMR and have an observation window extending out to the diffusional lifetime (∼0.5 ms in our geometry). The 1D NMR spectra are very rich in information because they offer clear limiting behavior for slow (>∼5 ms) and fast (<∼200 μs) chemical exchange as well as unique line shapes for intermediate exchange regimes. Furthermore, dynamic NMR experiments (as in Figure 3D) allow the acquisition of quantitative rate information about slow processes. Taken together, these collections of independent observables allow us to triangulate sensible parameter sets from which to build probable kinetic models. Again, the focus of this modeling is not to determine the most accurate values for every parameter of every variant but rather to assess the ability of the underlying scheme to encompass the range of observed behaviors and highlight the subtle differences among protein variants that cause them to occupy different regions of the parameter space.

Simulated spectra do indeed show that the diverse behavior among the proteins studied can be reproduced within different regions of the two-site model parameter space. The UV−vis, FCS, and NMR data from pH titrations of *ih*:GFP S65T, *s7*:GFP S65T, and *s7*+:GFP S65T exemplify these variations, and the simulated data are presented in Figure 7, with kinetic parameters listed in Table 1.

**Table 1**

| parameter | *ih*:GFP S65T (monomer) | *ih*:GFP S65T (dimer) | *s7*:GFP S65T | *s7*+:GFP S65T |
|---|---|---|---|---|
| $pK_{A'B}$ | 5.2 | 5.2 | 6.4 | 6.4 |
| $pK_{A'A}$ | 5.6 | 4.4 | 6.6 | 6.3 |
| $W$ | 0.1 | 1.0 | 4.0 | 0.0 |
| $k_{A'B}$ (s$^{-1}$) | 10000 | 100 | 50000 | 200 |
| $k_{A'A}$ (s$^{-1}$) | 10000 | 1000 | 5000 | 20000 |
| $k_{AB}$ (s$^{-1}$) | 8000 | 2000 | 500000 | 100000 |
| $k_{AB'}$ (s$^{-1}$) | 200 | 100 | 500 | 50 |
| $k_{BB'}$ (s$^{-1}$) | 200 | 1000 | 500 | 50 |

The two-site model is capable of replicating the full range of effects incompatible with the single-site framework. The contradictory chemical exchange time scales observed by FCS and NMR, particularly for *ih*:GFP S65T, can be reconciled under a two-site model and highlight the possibly misleading nature of either one taken separately because of their intrinsic sensitivity to chemical exchange processes on differing time scales. The simulations also reveal that phenomena in the NMR pH titrations such as disappearing peaks (e.g., Figure 3C and

Figure S4B of the Supporting Information) and peaks simultaneously showing both fast and slow exchange characteristics (e.g., Figure 4B) are natural consequences of certain parameter sets. Interestingly, we recently observed a similar case of tyrosine $^{13}$C NMR peaks vanishing during pH titration in a study of hydrogen bond networks in the keto-steroid isomerase active site.[39] Additionally, the UV−vis deviations from single-site titrations were neatly resolved through the influence of the site coupling parameter $W$. Small values of $W$ result in stretched isotherms characteristic of negative cooperativity, while large values of $W$ lead to strong internal buffering and titrations with mixed-state plateaus. In the following, we shift our attention toward the possible mechanistic implications of our observations when viewed through the lens of a two-state model.

**Mechanistic Implications.** The abnormal pH titration behaviors and heterogeneous kinetics of proton transfer in the GFP family have been noticed by others.[10,13] Specifically, Scharnagl et al. were the first to offer the explanation of electrostatic coupling between ionizable sites on the basis of simulations,[9] which was later empirically elaborated across many variants by Beltram and co-workers.[10] Further work with pH-jump studies revealed multistep kinetics that has been commonly interpreted as involving a PT to the protein surface followed by migration to the interior gated by conformational fluctuations, often implicating His148.[10,18−20,23] Our results are broadly consistent with this multisite ionization framework, and we find that the scheme in Figure 6 has good qualitative applicability across the mutants and circular permutants we investigated.

A handful of previous NMR studies on FPs may provide clues about the presence and nature of the long-lived species we report herein. Holak and co-workers have observed slow structural exchange processes occurring on millisecond time scales in two different contexts. In the first, they measured $^{19}$F NMR of fluoro-substituted tryptophan in CFP and observed two well-resolved chromophore fluorine peaks that were fit to a two-state chemical exchange model with a lifetime of 1.2−1.4 ms.[40] In the other, they observed twinning in a number of $^{1}$H−$^{15}$N HSQC peaks induced by the H148G mutation (on *s7*) in GFPuv, indicating two conformers in slow exchange.[24] These results highlight the unique structural role of strand 7 and its participation in slow dynamics.

Our experiments have indicated that the location of the circular permutation site with respect to *s7* has a dramatic effect on the observed ground-state proton transfer dynamics even while leaving sensitive photophysical properties such as quantum yield and ESPT largely unchanged (data not shown). In particular, CP at position 146 (*s7*:GFP S65T) significantly increases the rate of proton transfer as evidenced by the characteristic fast exchange NMR spectra obtained from pH titrations (Figure 4A).[b] Relocation of the CP by six residues upstream to position 140 (*s7*+:GFP S65T) completely alters the exchange dynamics on both long and short time scales and restores a long-lived B state (Figure 4B). We hypothesize that the CP at 146, just two residues upstream of His148 that is hydrogen bonded to the chromophore phenoxy, allows for greater flexibility and "fraying" of β-strand 7, thus accelerating the proton transfer processes.

Figure 6 details a kinetic model in which only one additional ionizable site, "X", is strongly coupled to the chromophore. In reality, there are probably many sites collectively modifying the chromophore proton affinity. UV−vis absorbance pH titration

curves, however, suggest that only a single site is responsible for most of the thermodynamic coupling. Several investigators[9,10] have speculated that Glu222 is playing this role because of both its proximity to the chromophore and its position as the terminal proton acceptor for ESPT. Moreover, the isosteric E222Q mutation results in pH titrations that are fit well by a simple single-site model (Figure S9 of the Supporting Information). Thus, as further mechanistic speculation, we will provisionally assume that "X" is Glu222.

The two most important thermodynamic parameters dictating the character of the titrations are the difference between the intrinsic $pK_a$ values ($pK_{A'A} - pK_{A'B}$), which determines the internal proton transfer equilibrium, and the coupling ($W$), which determines the degree of internal buffering. A property shared by the $int$PT proteins is a very high $W$. This factor causes them to have a very flat response to pH changes (dashed curves in Figure 2C). Furthermore, the high value for $W$ can be rationalized by recognizing that these proteins have a direct hydrogen bond network extending from the chromophore phenoxy to Glu222 over which ESPT occurs.[6,7] A direct consequence of this large coupling, $W$, is a broad plateau in the pH titration within which only intramolecular proton transfer occurs, making the fractional protonation of the chromophore solely determined by the difference in intrinsic $pK_a$ values (cf. $ih$:GFP S65S in Figure 2C and Figure S8 of the Supporting Information). Fortuitously, the circularly permuted $int$PT proteins differed in this internal PT equilibrium and allowed for the collection of NMR data in panels A and B of Figure 3 and thus the calculation of the basis spectra.

The $ext$PT proteins, in contrast, generally had much smaller values for $W$, which were manifested as titrations showing shallower slopes characteristic of negative cooperativity. All of these proteins had the S65T mutation, which has been shown in structures to disrupt the hydrogen bond network linking the chromophore and Glu222[7] and, according to our interpretation, should lead to a weakening of the coupling. The magnitude of the decrease in $W$ with S65T suggests that the majority of the coupling effect is mediated by specific hydrogen bonding rather than simply the through-space electrostatic repulsion that would be approximately the same for Ser65 or Thr65. It should be noted that the model still predicts efficient intramolecular GSPT between the chromophore and Glu222 despite the breakdown of ESPT. Interestingly, the one $ext$PT protein predicted to have a large $W$, $s7$:GFP S65T, is also the only one having appreciable fluorescence emission upon A-state excitation (data not shown). Further bolstering the case for Glu222 as the principal coupled ionizable site is the fact that upon mutation to glutamine (E222Q) the thermodynamic signature of any coupling is eliminated (see section S.9 of the Supporting Information for details).

The kinetic behavior of the $int$PT proteins is quite simple from the NMR perspective. As described above, the two-site model predicts that the only chemical exchange should be intramolecular proton transfer between A and B (see Figure S8 of the Supporting Information). The fact that a single population-weighted peak is observed in all $int$PT cases (see Figure 3A,B) implies that the internal proton shuttling between the chromophore and Glu222 is occurring rapidly. Early on, it was thought that the ground-state interconversion between the A state and B state was quite slow (hours).[5] Later, the work of Kennis et al. demonstrated a nanosecond scale ground-state return of the proton following ESPT (i.e., from the so-called "I

state" to the A state);[17] however, this does not directly report on the A- to B-state conversion time scale. Our findings clearly indicate on the basis of fast exchange on the NMR time scale that the A- to B-state exchange rate lies in the microsecond range or faster.

A recurrent theme from the global fit parameters for the $ext$PT proteins was the prominence of the doubly deprotonated state (B′ in Figure 6) in the slowest transition mode of the two-site exchange network. More specifically, the slowest mode of equilibrium chemical flux through the transition network was typically between B′ and a cluster comprised of B, A, and A′. This fact is manifested in the NMR spectra of most of the $ext$PT proteins by the chemical shift invariance of the 177 ppm peak. The kinetics of chemical exchange within the cluster {B, A, A′} are more variable and lead to qualitatively different outcomes. The pH-independent intramolecular PT connecting A and B is, by and large, the most rapid projected reaction. In cases of large $W$, namely $s7$:GFP S65T and the $int$PT proteins, this mixed-state {A, B} is identical to the pseudothermodynamic state termed "M" by Bizzarri et al.[10] In summary, the diverse NMR results can be largely understood as arising from B′ in slow exchange with the rapidly interconverting {A, B} that is, in turn, exchanging with A′ with a variable rate among protein variants. Below, we briefly consider the consequences of this interpretation on the proteins simulated in Figure 7.

For both the dimer and monomer forms of $ih$:GFP S65T, the interconversion between {A, B} and A′ falls within an intermediate exchange regime. This leads to the unusual situation in which the chromophore peak in the NMR spectrum seems to disappear almost entirely (see Figure 3C and Figure S4B of the Supporting Information). The simulated spectra (Figure 7C,F) suggest that the apparent loss of intensity is, in fact, due to slow exchange of the well-defined B′ peak with an extremely broad peak comprised of {A, B} in exchange with A′ near the coalescence point. As the pH is further lowered, the signal reconsolidates toward a pure A′ state.

The lack of a defined B′ peak in $s7$:GFP S65T (Figure 4A) is caused by a strong site coupling term ($W$), which causes the apparent high-pH limit to be the {A, B}, or "M", with a negligible population of B′ (Figure 7G). Thus, the single peak with a changing chemical shift is indicative of a fast exchange process between {A, B} and A′.

The situation for $s7+$:GFP S65T (Figure 4B) shares certain features with the two preceding cases. Like $ih$:GFP S65T, it has a well-resolved B′ peak in exchange with some combination of B, A, and A′. The kinetic behavior of this cluster, however, is more similar to that of $s7$:GFP S65T in which the three states are in mutually fast exchange. Together, these factors lead to the two apparent populations in the NMR spectra of B′ at the deprotonated basis and {B, A, A′}, whose composition changes with a decrease in pH. The saturation transfer experiments described in the Results and detailed in section S.6 of the Supporting Information demonstrate that the two populations in the $s7+$:GFP S65T titration (Figure 4B) are in fact in chemical exchange on a time scale of tens of milliseconds and thus consistent with the two-site kinetic scheme.

One of the most striking features from the data is the very long lifetime of the doubly deprotonated state B′. The physical basis for the slowness of this transition out of B′ is not clear from our measurements. The prolonged time scale of the proton transfer is suggestive of an accompanying gross structural change and could be related to the $s7$ conformations observed in previous NMR studies.[24] Krishnamoorthy and co-

workers also suggested that PT is being gated by large conformational changes on the basis of a significant viscosity dependence to pH-jump fluorescence kinetics.[18] Our results indicate that a second ionizable site is playing a large kinetic role and could potentially be modulating the rate of these conformational fluctuations. Both the chromophore phenoxy and Glu222 participate in a hydrogen bond network that extends to s7, and it is possible that proton transfers are actuating larger structural rearrangements.

The work presented here highlights the array of factors influencing the GFP chromophore protonation and deprotonation kinetics. Most prominent among these seems to be the influence of the secondary ionizable site, Glu222, primarily via its influence on hydrogen bond networks encompassing the chromophore. Interestingly, subtle and distal changes due to protein circular permutation can profoundly alter the proton transfer kinetics while still qualitatively fitting within the coupled two-site framework. This implies that efforts to approach this problem computationally would likely require simulation of the complete structure using quantum mechanics/molecular mechanics methods with a quantum region large enough to capture the proton transfer chemical reactions. The size of the protein (238 residues), the time scales of the proton transfers (up to tens of milliseconds), and the requirement for reactive dynamics all make this a daunting challenge. The GFP system could serve as a prototype for these types of advanced calculations whose predictions may be compared against the rich variety of specific experimental observables that can be obtained because of the spectral and chemical uniqueness of the chromophore.

## CONCLUSION

We have introduced a site-specific $^{13}$C NMR probe to provide a nonoptical observable for proton transfer in GFP. The union of this technique with the optical methods of UV−vis absorbance and FCS allowed for a comprehensive view of the multisite nature of the system. Our data from this approach underscore the potential pitfalls of relying on single kinetic measurement techniques because of the bias introduced by their intrinsic observational time scale. The kinetics of proton transfer in Superfolder GFP are shown to correspond well to a model with two thermodynamically coupled titratable sites across a range of mutants and circular permutants. The PT rates are significantly impacted by both mutational disruptions to internal hydrogen bond networks and the presumed structural rigidity of $\beta$-strand 7 as probed by circular permutation.

## ASSOCIATED CONTENT

### Ⓢ Supporting Information

Complete sequences for all protein constructs, concentration- and light-dependent effects on PT dynamics, NMR saturation transfer for connectivity analysis, fitting and simulation methods for FCS and NMR, E222Q titrations, and local His tag effects. This material is available free of charge via the Internet at http://pubs.acs.org.

## AUTHOR INFORMATION

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

## ABBREVIATIONS

GFP, green fluorescent protein; PT, proton transfer; GSPT, ground-state proton transfer; ESPT, excited-state proton transfer; NMR, nuclear magnetic resonance; FCS, fluorescence correlation spectroscopy; ACF, autocorrelation function; CP, circular permutation.

## ADDITIONAL NOTES

[a]In principle, the data are convolved with the $T_1$ relaxation time; however, that time [∼5.7 s (data not shown)] was found to be several orders of magnitude longer than the chemical exchange lifetime.
[b]It is interesting to note that CPs or insertions into this area of the protein have found many successful applications in chemical sensing with recognition domains.[41]

## REFERENCES

(1) Dedecker, P., De Schryver, F. C., and Hofkens, J. (2013) Fluorescent Proteins: Shine on, You Crazy Diamond. *J. Am. Chem. Soc.* 135, 2387−2402.

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

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

(4) Lukyanov, K. A., Dmitry, M. C., Sergey, L., and Vladislav, V. V. (2005) Photoactivatable fluorescent proteins. *Nat. Rev. Mol. Cell Biol.* 6, 885−890.

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

(6) 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.

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

(8) 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.

(9) Scharnagl, C., Raupp-Kossmann, R., and Fischer, S. F. (1999) Molecular basis for pH sensitivity and proton transfer in green fluorescent protein: Protonation and conformational substates from electrostatic calculations. *Biophys. J.* 77, 1839−1857.

(10) Bizzarri, R., Nifosi, R., Abbruzzetti, S., Rocchia, W., Guidi, S., Arosio, D., Garau, G., Campanini, B., Grandi, E., Ricci, F., Viappiani,

C., and Beltram, F. (2007) Green fluorescent protein ground states: The influence of a second protonation site near the chromophore. *Biochemistry 46*, 5494−5504.

(11) Fron, E., Flors, C., Schweitzer, G., Habuchi, S., Mizuno, H., Ando, R., De Schryver, F. C., Miyawaki, A., and Hofkens, J. (2007) Ultrafast excited-state dynamics of the photoswitchable protein dronpa. *J. Am. Chem. Soc. 129*, 4870−4871.

(12) Lukacs, A., Haigney, A., Brust, R., Addison, K., Towrie, M., Greetham, G. M., Jones, G. A., Miyawaki, A., Tonge, P. J., and Meech, S. R. (2013) Protein Photochromism Observed by Ultrafast Vibrational Spectroscopy. *J. Phys. Chem. B 117*, 11954−11959.

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

(14) Cotlet, M., Goodwin, P. M., Waldo, G. S., and Werner, J. H. (2006) A comparison of the fluorescence dynamics of single molecules of a green fluorescent protein: One- versus two-photon excitation. *ChemPhysChem 7*, 250−260.

(15) Liu, Y., Kim, H. R., and Heikal, A. A. (2006) Structural basis of fluorescence fluctuation dynamics of green fluorescent proteins in acidic environments. *J. Phys. Chem. B 110*, 24138−24146.

(16) Bosisio, C., Quercioli, V., Collini, M., D'Alfonso, L., Baldini, G., Bettati, S., Campanini, B., Raboni, S., and Chirico, G. (2008) Protonation and conformational dynamics of GFP mutants by two-photon excitation fluorescence correlation spectroscopy. *J. Phys. Chem. B 112*, 8806−8814.

(17) Kennis, J. T. M., Larsen, D. S., van Stokkum, I. H. M., Vengris, M., van Thor, J. J., and van Grondelle, R. (2004) Uncovering the hidden ground state of green fluorescent protein. *Proc. Natl. Acad. Sci. U.S.A. 101*, 17988−17993.

(18) Mallik, R., Udgaonkar, J. B., and Krishnamoorthy, G. (2003) Kinetics of proton transfer in a green fluorescent protein: A laser-induced pH jump study. *Proc.—Indian Acad. Sci., Chem. Sci. 115*, 307−317.

(19) Saxena, A. M., Udgaonkar, J. B., and Krishnamoorthy, G. (2005) Protein dynamics control proton transfer from bulk solvent to protein interior: A case study with a green fluorescent protein. *Protein Sci. 14*, 1787−1799.

(20) Abbruzzetti, S., Grandi, E., Viappiani, C., Bologna, S., Campanini, B., Raboni, S., Bettati, S., and Mozzarelli, A. (2005) Kinetics of acid-induced spectral changes in the GFPmut2 chromophore. *J. Am. Chem. Soc. 127*, 626−635.

(21) McAnaney, T. B., Zeng, W., Doe, C. F., Bhanji, N., Wakelin, S., Pearson, D. S., Abbyad, P., Shi, X., Boxer, S. G., and Bagshaw, C. R. (2005) Protonation, photobleaching, and photoactivation of yellow fluorescent protein (YFP 10C): A unifying mechanism. *Biochemistry 44*, 5510−5524.

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

(23) Agmon, N. (2005) Proton pathways in green fluorescence protein. *Biophys. J. 88*, 2452−2461.

(24) 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.

(25) Baturin, S. J., Okon, M., and McIntosh, L. P. (2011) Structure, dynamics, and ionization equilibria of the tyrosine residues in *Bacillus circulans* xylanase. *J. Biomol. NMR 51*, 379−394.

(26) Oktaviani, N. A., Pool, T. J., Kamikubo, H., Slager, J., Scheek, R. M., Kataoka, M., and Mulder, F. A. A. (2012) Comprehensive Determination of Protein Tyrosine p$K_a$ Values for Photoactive Yellow Protein Using Indirect C-13 NMR Spectroscopy. *Biophys. J. 102*, 579−586.

(27) Takeda, M., Miyanoiri, Y., Terauchi, T., Yang, C.-J., and Kainosho, M. (2014) Use of H/D isotope effects to gather information about hydrogen bonding and hydrogen exchange rates. *J. Magn. Reson. 241*, 148−154.

(28) Fafarman, A. T., Sigala, P. A., Schwans, J. P., Fenn, T. D., Herschlag, D., and Boxer, S. G. (2012) Quantitative, directional measurement of electric field heterogeneity in the active site of ketosteroid isomerase. *Proc. Natl. Acad. Sci. U.S.A. 109*, E299−E308.

(29) Takeda, M., Jee, J., Ono, A. M., Terauchi, T., and Kainosho, M. (2009) Hydrogen Exchange Rate of Tyrosine Hydroxyl Groups in Proteins As Studied by the Deuterium Isotope Effect on $C_\zeta$ Chemical Shifts. *J. Am. Chem. Soc. 131*, 18556−18562.

(30) 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.

(31) Robinson, G., Kuchel, P. W., Chapman, B. E., Doddrell, D. M., and Irving, M. G. (1985) A Simple Procedure for Selective Inversion of NMR Resonances for Spin Transfer Enzyme Kinetic Measurements. *J. Magn. Reson. 63*, 314−319.

(32) Kapusta, P. (2010) *Absolute Diffusion Coefficients: Compilation of Reference Data for FCS Calibration*, PicoQuant, GmbH, Berlin.

(33) Laurence, T. A., Fore, S., and Huser, T. (2006) Fast, flexible algorithm for calculating photon correlations. *Opt. Lett. 31*, 829−831.

(34) Kent, K. P., Oltrogge, L. M., and Boxer, S. G. (2009) Synthetic control of green fluorescent protein. *J. Am. Chem. Soc. 131*, 15988−15989.

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

(36) Xiang, H., Bell, A. F., and Tonge, P. J. (2002) Isotopic labeling and normal-mode analysis of a model green fluorescent protein chromophore. *J. Phys. Chem. B 106*, 6056−6066.

(37) Mizuno, H., Mal, T. K., Walchli, M., Fukano, T., Ikura, M., and Miyawaki, A. (2010) Molecular basis of photochromism of a fluorescent protein revealed by direct $^{13}$C detection under laser illumination. *J. Biomol. NMR 48*, 237−246.

(38) Efron, B., and Tibshirani, R. (1986) Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy. *Statistical Science*, 54−75.

(39) Sigala, P. A., Fafarman, A. T., Schwans, J. P., Fried, S. D., Fenn, T. D., Caaveiro, J. M. M., Pybus, B., Ringe, D., Petsko, G. A., Boxer, S. G., and Herschlag, D. (2013) Quantitative dissection of hydrogen bond-mediated proton transfer in the ketosteroid isomerase active site. *Proc. Natl. Acad. Sci. U.S.A. 110*, E2552−E2561.

(40) Seifert, M. H., Ksiazek, D., Azim, M. K., Smialowski, P., Budisa, N., and Holak, T. A. (2002) Slow exchange in the chromophore of a green fluorescent protein variant. *J. Am. Chem. Soc. 124*, 7932−7942.

(41) Nakai, J., Ohkura, M., and Imoto, K. (2001) A high signal-to-noise $Ca^{2+}$ probe composed of a single green fluorescent protein. *Nat. Biotechnol. 19*, 137−141.

Supporting Information for:

# "Ground-state proton transfer kinetics in Green Fluorescent Protein (GFP)"

*Luke M. Oltrogge, Quan Wang, 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 the sequence of Superfolder GFP (sfGFP) which was used for all experiments due to its robust stability, high expression yield, and tolerance to circular permutation.  These proteins differ from wtGFP by about a dozen mutations selected largely for improved folding properties, and they do not appreciably affect the fluorescence properties which remain nearly identical to wild-type.[1]  The circular permutants have all had the native N- and C-termini fused with the linker sequence, GGTGGS.  These constructs were designed by us and synthetically prepared by Genscript.

In all cases below the residue triad forming the chromophore (on *ih*) is highlighted in green, His148 (on *s7*) is highlighted in yellow, the original terminal linker is highlighted in cyan, and mutation locations are indicated with gray highlight and superscripts.


*ih*:GFP

MGHHHHHHSSGGKLPVPWPTLVTTLS[a]YGVQCFSRYGTRGSGSIEGRHSGSGSPDHMKR
HDFFKSAMPEGYVQERTISFKDDGKYKTRAVVKFEGDTLVNRIELKGTDFKEDGNILGH
KLEYNFNSHNVYITADKQKNGIKANFTVRHNVEDGSVQLADHYQQNTPIGDGPVLLPD
NHYLSTQTV[b]LSKDPNEKRDHMVLLE[c]FVTAAGITHGMDELYGGTGGSASQGEELFTGV
VPILVELDGDVNGHKFSVRGEGEGDATIGKLTLKFISTT

**a.**  S65T, **b.**  V206K, **c.**  E222Q


*s7*:GFP

MGHHHHHHSSGGNSHNVYITADKQKNGIKANFTVRHNVEDGSVQLADHYQQNTPIGDG
PVLLPDNHYLSTQTV[a]LSKDPNEKRDHMVLLE[b]FVTAAGITHGMDELYGGTGGSASQGE
ELFTGVVPILVELDGDVNGHKFSVRGEGEGDATIGKLTLKFISTTGKLPVPWPTLVTTLS[c]
YGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTISFKDDGKYKTRAVVKFEGDTLVNRI
ELKGTDFKEDGNILGHKLEYNF

**a.**  V206K, **b.**  E222Q, **c.**  S65T

*s7+*:GFP

MGHHHHHHRSSGKLEYNFNSHNVYITADKQKNGIKANFTVRHNVEDGSVQLADHYQQ
NTPIGDGPVLLPDNHYLSTQTKLSKDPNEKRDHMVLLE<sup>a</sup>FVTAAGITHGMDELYGGTGG
SASQGEELFTGVVPILVELDGDVNGHKFSVRGEGEGDATIGKLTLKFISTTGKLPVPWPT
LVTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTISFKDDGKYKTRAVVKFEGD
TLVNRIELKGTDFKEDGNILGH

**a.** E222Q

*s10*:GFP

MGSSHHHHHHSSGLVPGGSHMLPDNHYLSTQTVLSKDPNEKRDHMVLHEYVNAAGIT
HGMDELYGGTGGSASQGEELFTGVVPILVELDGDVNGHKFSVRGEGEGDATIGKLTLKF
ISTTGKLPVPWPTLVTTLSYGVQAFSRYPDHMKRHDFFKSAMPEGYVQERTISFKDDGK
YKTRAVVKFEGDTLVNRIELKGTDFKEDGNILGHKLEYNFNSHNVYITADKQKNGIKAN
FTVRHNVEDGSVQLADHYQQNTPIGDGPVL

*s1*:GFP *(un-permuted)*

MGSSHHHHHHSSGLVPRGSHMGGTSSKGEELFTGVVPILVELDGDVNGHKFSVRGEGE
GDATIGKLTLKFICTTGKLPVPWPTLVTTLSYGVQCFSRYPDHMKRHDFFKSAMPEGYV
QERTISFKDDGKYKTRAVVKFEGDTLVNRIELKGTDFKEDGNILGHKLEYNFNSHNVYI
TADKQKNGIKANFTVRHNVEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQTVLSKD
PNEKGTRDHMVLHEYVNAAGIT

### S.2    Determination of Protonation Equilibrium by UV/Vis Absorbance

In order to estimate the fractional ionization of the chromophore at arbitrary pH it was necessary to establish the intrinsic extinction coefficients for the protonated and deprotonated basis states. Throughout this text we make the assumption that both neutral and both ionized forms (A, A′ and B, B′ from Fig. 6) have identical absorbance spectra, and, moreover, that there is no significant pH-dependence to either basis state. This approximation of only two optical states is well supported by the generally clean isosbestic points (*e.g.* Fig. 2B). The deviations that do exist are primarily due to partial denaturation at the extrema of pH.

In general the spectrum may be expressed as a linear combination of basis spectra using the Beer-Lambert law as,

(Eq. S1) $$S(\lambda) = bc\left[(1 - p_B)S_A(\lambda) + p_B S_B(\lambda)\right]$$

where $b$ is the pathlength in cm, $c$ is the molar concentration, $p_B$ is the fraction ionized, and $S_A(\lambda)$ and $S_B(\lambda)$ are the basis spectra for the A and B states in units of $M^{-1}cm^{-1}$.

The basis spectra can be obtained in a straightforward way for the model titratable protein *ih*:GFP S65T. This protein features a pH titration well approximated by a single-site model giving an apparent $pK_a$ of ~5.9. Consequently, at pH 8.0 and higher the protein has >99% B-state and thus directly provides $S_B(\lambda)$. Additionally, the A-state has negligible absorbance at the B-state maximum (~485 nm) so $p_B$ may be accurately estimated within the pH titration from this relative absorbance. The integrated fluorescence emission from 485 nm excitation follows exactly the same pH trend as the absorbance and further confirms the absence of A-state absorbance at this wavelength. Together this information enables the calculation of the A-state basis spectrum, $S_A(\lambda)$. The scaling of $S_A(\lambda)$ and $S_B(\lambda)$ to extinction coefficients was performed

using the method of Ward *et al.* with 0.1 M NaOH denaturation and the known extinction

coefficient of the denatured deprotonated chromophore of 44,100 $M^{-1}cm^{-1}$ at 447 nm.[2]

In practice, the concentration ($c$) and fraction deprotonated ($p_B$) were often determined

using a two point calculation from the absorbances at the A-state and B-state maxima (~395 nm

and ~485 nm, respectively).

(Table S1)

| | $\varepsilon_{395}$ ($M^{-1}cm^{-1}$) | $\varepsilon_{485}$ ($M^{-1}cm^{-1}$) |
|---|---|---|
| *A-state* | 30,700 | 0 |
| *B-state* | 4,600 | 57,200 |

After obtaining an experimental spectrum, $S(\lambda)$, $c$ and $p_B$ were calculated by simultaneous

solution of Eq. S1 at 395 nm and 485 nm. The analytical equations for this procedure are,

(Eq. S2)
$$c = \frac{S(485)[S_A(395) - S_B(395)] + S(395)[S_B(485) - S_A(395)]}{S_A(395)S_B(485) - S_A(485)S_B(395)}$$

(Eq. S3)
$$p_B = \frac{S(485)S_A(395) - S(395)S_A(485)}{S(485)[S_A(395) - S_B(395)] + S(395)[S_B(485) - S_A(395)]}$$

### S.3     FCS Experiments and Fitting

The autocorrelation functions (ACFs) were fit to a model containing three decay components: 1) the diffusional lifetime, 2) the light-dependent dark state formation—likely a triplet state, and 3) the pH-dependent protonation/deprotonation chemical exchange, which is the quantity we seek.

It was experimentally determined through measurements of a sub-diffraction sized fluorescent bead translated through the focal volume on a piezo-electric stage that the detection volume was well approximated by a Gaussian ellipsoid point spread function with $r_0 \sim 0.4\mu m$ and $z_0 \sim 1.6\mu m$. With this standard assumption for the focal volume the complete expression for the autocorrelation function (ACF) is,

$$(\text{Eq. S4}) \qquad ACF(t) \propto \left( 1 - F_{dark} + F_{dark}e^{\frac{-t}{\tau_{dark}}} \right) * \left( 1 - F_{prot} + F_{prot}e^{\frac{-t}{\tau_C}} \right) * \left( \frac{1}{1 + \frac{4Dt}{r_0^2}} \right) * \left( \frac{1}{1 + \frac{4Dt}{z_0^2}} \right)^{\frac{1}{2}}$$

where $F_{dark}$ is the fraction of the population driven into a dark state with lifetime $\tau_{dark}$ at a given light intensity, $F_{prot}$ is the equilibrium fraction of the population in the dark protonated state with apparent chemical exchange lifetime $\tau_C$, $D$ is the diffusion coefficient, and $r_0$ and $z_0$ describe the focal volume dimensions. For a single-site titration model the expression for $\tau_C$ can be written as a function of pH and the rate constants,

$$(\text{Eq. S5}) \qquad \tau_C(pH) = \frac{1}{k_{deprot} + k_{prot}10^{-pH}}$$

The protein samples were measured for a range of different pH values. Most trajectories were recorded with an irradiance of 10kW/cm$^2$, however, some were also measured with 2.5kW/cm$^2$ in order to assess the effect of the light intensity. It was observed that the pH-dependent flickering was unaffected by this change but that the short timescale decay was lessened in magnitude but unchanged in rate. This evidence is consistent with its tentative assignment to a dark triplet state.

In practice, the fitting was performed by first measuring the sample under high pH conditions in which the chromophore is always deprotonated and using only the diffusional and light-dependent triplet parts of the calculated ACF in the fit. The data were normalized at the ACF plateau occurring at one microsecond. We then assumed that the protein retains a similar diffusion coefficient and photophysical properties and fixed these parameters while varying only the chemical exchange parameters for the samples at all other pH values. Using the expression for $\tau_C$ from Eq. S5 we used $k_{deprot}$ and $k_{prot}$ as global fit parameters for a non-linear least squares optimization in Matlab.

## S.4    $^{13}C_\zeta$-Tyrosine Spectrum and Normalization Procedure

All of the protein variants contained nine tyrosine residues. The $^{13}C_\zeta$ label was incorporated into all of these as confirmed by a shift of +9 Da observed in intact protein electrospray ionization mass spectrometry (data not shown). In small unstructured peptides the $^{13}C_\zeta$ resonances of Tyr have been measured to occur at 155.5 ppm in the neutral form and 166.3 ppm in the ionized form.[3] The cluster of peaks between 154 and 158 ppm shown in Fig. S1 arises from the eight protonated tyrosines in the protein. The ninth tyrosine is Tyr66 which is incorporated into the chromophore has its chemical shift increased to 162.0 ppm in the protonated form and 177.1 ppm in the deprotonated form by virtue of the extended electronic conjugation. The tyrosine cluster showed only a small dependence on pH through the titration series due the relatively high $pK_a$ of all tyrosines ( ~10).[4]

All NMR spectra were normalized by the same method using the protonated tyrosine peaks. The integrated area underneath these peaks was normalized to eight. Our choice to use the tyrosine $^{13}C_\zeta$ peaks for normalization—as opposed to some other feature such as the aliphatic region—was based on the large intensity due to these peaks (see Fig. S2) and the fact that their abundance should be directly related to the chromophore peak intensity arising from the incorporation of the same label and thus insensitive to variations in labeling efficiency. That the backbone carbonyl resonance bands overlap nicely in Fig. 3 A&C and Fig. 4 B&C reflects the fact that the isotope labeling was reproducible and essentially complete—in accordance with the mass spectrometry results.
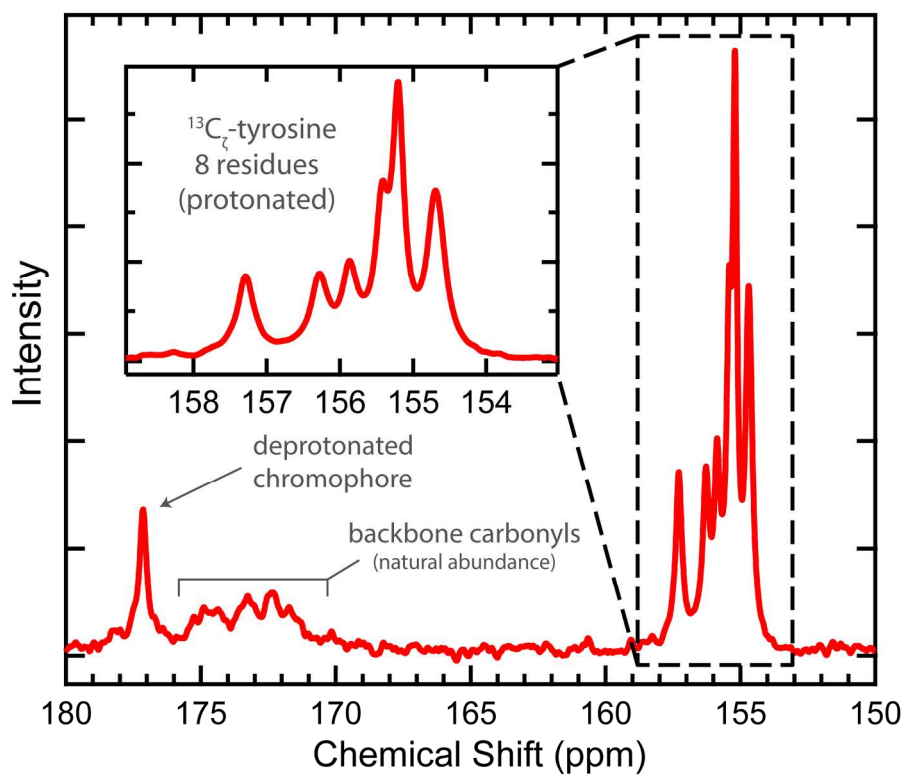
*Fig. S1*

1D $^{13}$C-NMR spectrum of *ih*:GFP S65T at pH 8.0.  The spectrum is composed of approximately 8,000 transients collected on a 300MHz Varian spectrometer with a 10 Hz exponential line-broadening function.
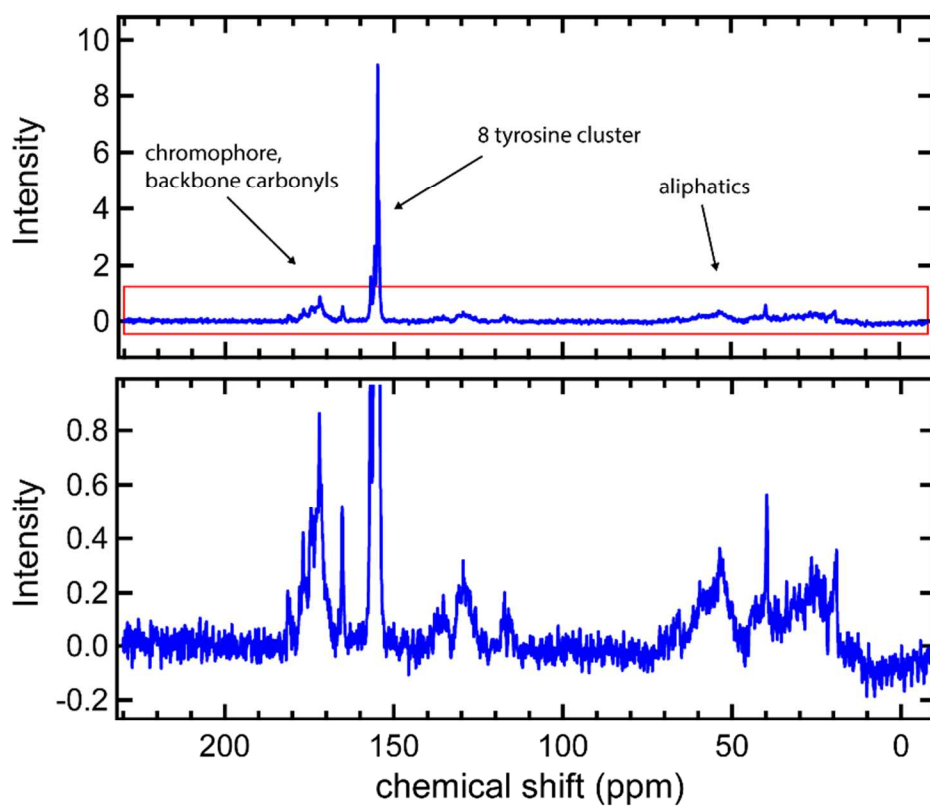
*Fig. S2*

Complete 1D $^{13}$C-NMR spectrum of a representative protein sample (*s7+*:GFP S65T at pH 6.6) which highlights the relative intensity of the labeled peaks to the peaks arising from $^{13}$C at natural abundance from the rest of the protein.
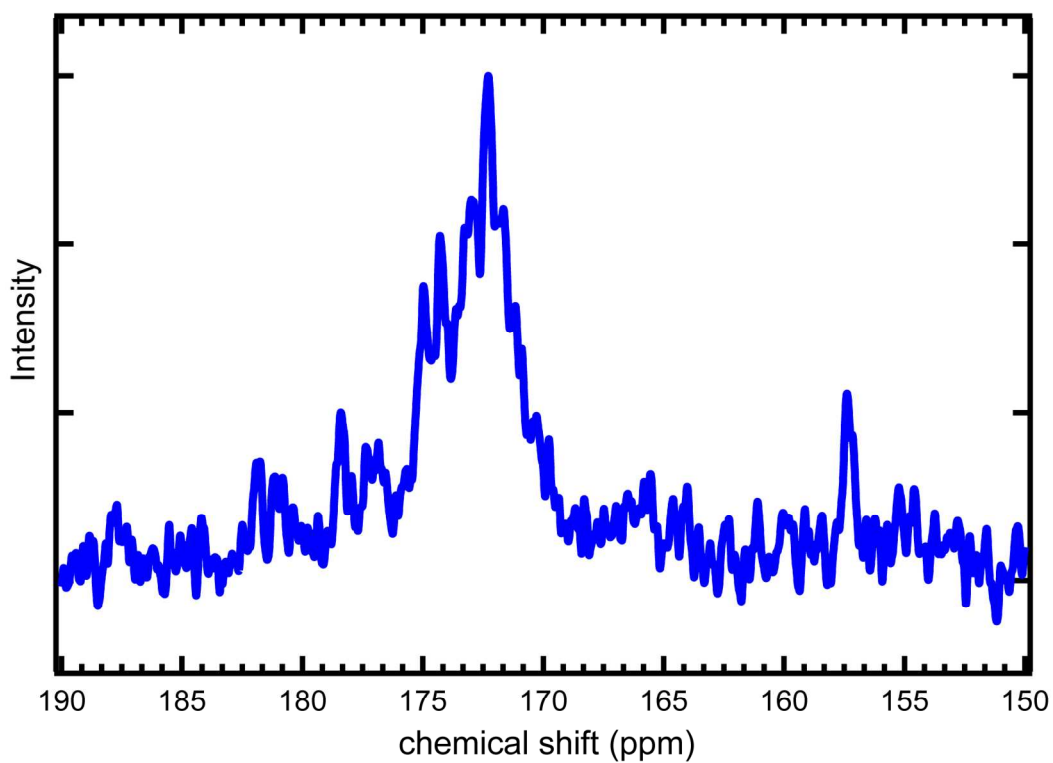
*Fig. S3*

1D ^{13}C-NMR spectrum of unlabeled protein in the region of interest. The broad cluster of peaks centered around 173 ppm is due to backbone and side-chain carbonyl carbons at natural abundance. The feature at 158 ppm is likely due to the arginine $^{13}C_{\zeta}$ with its sharpness due to the relatively free rotation of the solvent exposed side-chain.[5]

## S.5    Controls for NMR and FCS Artifacts

Beyond the quantitative disagreement with the kinetic selective inversion experimental results, the qualitative appearance alone of the low pH $^{13}$C-NMR spectra—in which the B-state peak decreases in intensity without changing position—precludes rates faster than ~1ms from lineshape predictions (Fig. 3C).  Seeing the striking difference between NMR and FCS we sought to investigate the possible artifacts arising from either technique to explain the discrepancy.  The two most significant differences between the experimental conditions are the concentration (~500pM for FCS and 1-2mM for NMR) and the illumination (~10kW/cm$^2$ for FCS and 0W/cm$^2$ for NMR).  Many FPs are known to oligomerize at high concentration with wtGFP having a dimerization constant reported as ~100μM.[6]  It is possible that sfGFP is forming dimers at the high NMR concentrations and changing the structural dynamics of the protein or even occluding the proton entry point.  On the other hand, in FCS the protein is subject to very intense illumination not present in the NMR.  Perhaps this light is driving some saturating behavior in which the proton transfer dynamics are altered.

### *S.5.a    Concentration Effects*

We tested the effect of concentration by creating limiting situations: first, by guaranteeing monomers in the NMR measurement and second, by engineering dimers for an FCS measurement.

### *NMR of Monomers*

It was noted very early by Ward and coworkers that the absorbance spectrum of wtGFP depended on the concentration in a way that caused a shift towards greater A-state absorbance.[2]

We measured the absorbance spectra of *ih*:GFP S65T up to concentrations as high as 1mM, taking care with the pathlength to avoid saturation and maintain accurate readings. Indeed, we also see a concentration-dependent increase in the A-state absorbance that suggests a dimerization constant around 100μM (see Fig. S4A, green). The mutation V206K, first reported by Tsien and coworkers,[7] introduces a positive charge in the middle of the hydrophobic dimer interface on strand 10 pointing outward into the solution; it was measured to change the dimerization constant from ~100μM to ~3mM by ultracentrifugation. This mutation was introduced to make *ih*:GFP S65T, V206K and the absorbance experiment was repeated. The resulting absorption spectra were unchanged up to at least 1mM thereby suggesting that V206K is in fact preventing dimerization (see Fig. S4A, blue).

The 1D $^{13}$C-NMR spectrum of the monomeric *ih*:GFP S65T, V206K labeled with $^{13}$C$_\zeta$-tyrosine was measured at pH 6.0 and 200μM on a 500MHz Bruker spectrometer with an indirect detection cryoprobe. The B-state peak stayed at exactly the same chemical shift indicating that indeed it remained in a slow exchange regime on the NMR timescale (Fig. S4B). At concentrations appreciably higher than 100μM *ih*:GFP S65T is effectively a dimer and, as a result, has a chromophore $pK_a$ which is perturbed upward. This fact allowed us to observe higher A-state fractions for any given pH while avoiding the problems of protein precipitation encountered at pH less than 6.0. For this reason a complete NMR titration and the time-resolved NMR experiments were only practical using the protein not bearing the V206K mutation. Consequently the NMR results from Fig. 3C&D are more accurately described as representing the proton dynamics of the dimer. Nonetheless, the NMR spectra in Fig. S4B establish even in the strictly monomeric protein that there is a long-lived deprotonated species.

*FCS of Dimers*

Nearly all x-ray structures of GFPs show a hydrophobic protein interface centered at residue 206 on strand 10.[6] A covalent dimer version of the protein was created with the use of a disulfide bridge across the native dimer interface. We prepared the mutant *ih*:GFP S65T, V206C and allowed air oxidation of the sulfhydryl before size exclusion purification of the dimer which was confirmed pure by mass spectrometry and SDS-PAGE. FCS measurements were made for the complete pH titration as for the other samples. The diffusion coefficient resulting from the fit in the high pH limit was approximately 80% of the value for the monomer. This is in close agreement with the expectation from the Stokes-Einstein equation if we assume a spherical shape and a doubling of the hydrodynamic volume.

The pH-dependent flickering rates were somewhat slower than that observed in the monomer but not nearly enough to bring it into agreement with the NMR results in Fig. 3C&D. It is interesting to note that the dimer ACFs in Fig. 5B fit very poorly to a single-site model and qualitatively imply negative cooperativity (fits not shown). This result is further evidence for the multi-site character of the titration.

In summary the experiments above led us to conclude that the *ih*:GFP S65T monomer was still undergoing a slow exchange process that led to a particularly long-lived deprotonated species. The FCS measurements on an engineered covalent dimer of *ih*:GFP S65T revealed dynamics that were slightly slower but still much faster than the apparent rate from NMR. This small reduction in rate may be due to partial occlusion of the putative proton entry site through strand 7 which is immediately adjacent to strand 10 which lies at the center of the homodimer interface.
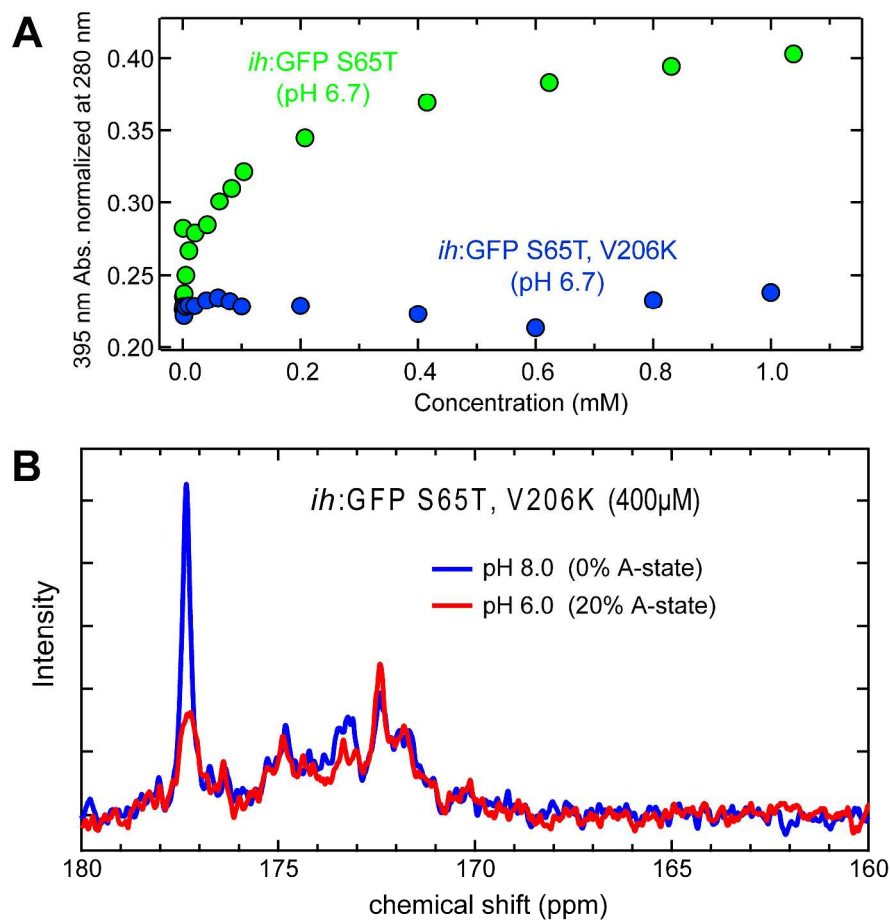
*Fig. S4*

**A)** Concentration dependence of A-state absorbance normalized to the absorbance at 280nm. The apparent $pK_a$ of the dimer is higher than the monomer and so this relative absorbance is a proxy for the extent of dimerization. *ih*:GFP S65T (green) shows a clear dimerization isotherm with an apparent $K_D$ of ~100μM. Upon addition of an aggregation-suppressing mutation *ih*:GFP S65T, V206K (blue) shows no sign of dimerization up to 1mM. **B)** 1D $^{13}$C-NMR spectra of *ih*:GFP S65T, V206K at 400μM and pH 8.0 and 6.0. Under these conditions all the protein is monomeric. The fact that the peak at 177.2 ppm does not shift with pH implies that rate of proton transfer is still in the slow exchange regime.

**S.5.b** *Light Effects*

Fluorescent proteins are known to undergo light-dependent processes on many timescales from ESPT (ps) to triplet state formation (µs) to blinking and photobleaching (s). As mentioned above in *FCS Experiments and Fitting* (§S.3), we tested the sensitivity of the ACFs to light intensity by measuring fluorescent trajectories at 10 and 2.5 $kW/cm^2$ and the only effect was on the magnitude of the presumed triplet state formation. It is still possible that, even at this lower intensity, we are still under sufficiently high irradiance to be at saturation for some other light-dependent process with consequences for PT kinetics. To control for this possibility we conducted a 1D $^{13}$C-NMR experiment under laser illumination of the B-state absorbance.

*In Situ Illumination 1D $^{13}$C-NMR*

The *in situ* illumination inside the magnet was done in a similar fashion to Mizuno *et al.*[8] Briefly, a 473nm diode-pumped solid state laser was coupled into a multimode optical fiber which was run into a 10-mm Shigemi tube (Fig. S5A). The plunger portion of the tube was partially filled with water for better refractive index matching with the glass. To mitigate the inner-filter effect as much as possible the protein concentration was lowered to 400µM and the plunger was set to create a disk of sample 4mm thick. The final output power from the optical fiber into the sample was 105mW. In order to minimize photobleaching over the 12 hour experiment, a shutter was triggered by the auxiliary NMR pulse output to open 100 ms before and then close after the acquisition time (0.5 s) for the duration of the interpulse delay interval (Fig. S5C).

1D $^{13}$C-NMR measurements of *ih*:GFP S65T at pH 6.6 were made with this setup first for 12 hrs. in the dark and then 12 hrs. with irradiation. It was found that the sample suffered less

than 2% photobleaching over the acquisition as determined by UV/Vis absorbance (Fig. S5B). A comparison of the dark and illuminated NMR spectra revealed that they were identical within the noise (Fig. S5D). More specifically, the peak at the B-state basis (177ppm) did not shift which indicates that it remains in a slow exchange regime even under illumination. In the FCS experiment lowering the power had no effect on the ACF aside from changing the dark triplet state component. This suggests that the light intensity is far within a saturating regime of some unknown light-driven process if it is playing a role at all. The fact that the significantly dimmer NMR experiment also experiences no change is consistent with the hypothesis that the proton transfer kinetics are unaffected by illumination.
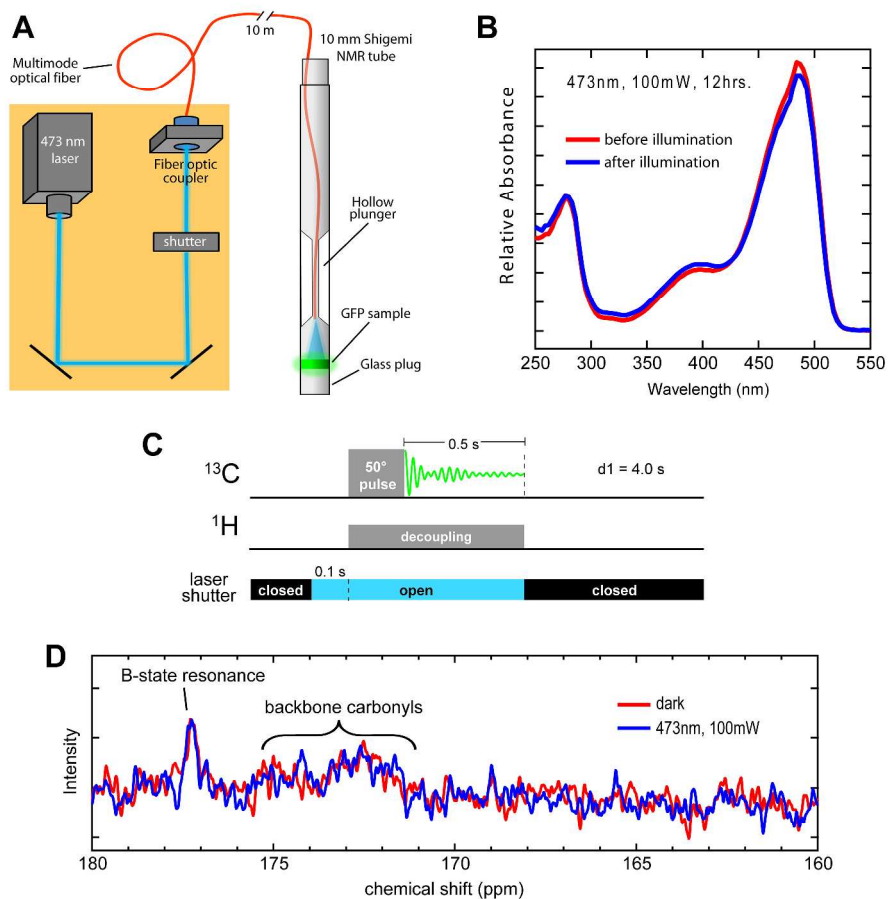
*Fig. S5*

**A)** Schematic drawing of the laser-illumination setup (drawing not to scale).  **B)** Absorbance spectra of *ih*:GFP S65T before and after NMR measurement.  The protein suffers about 2% photobleaching over the course of the experiment.  **C)** Diagram of pulse sequence used in illumination experiments.  **D)** 1D $^{13}$C-NMR spectra of *ih*:GFP S65T for identical experiments in the dark (red) and with 100mW, 473 nm laser illumination (blue).  The spectra are indistinguishable from each other within the noise and imply that the light has no effect of proton transfer kinetics.

## S.6    NMR Saturation Transfer

To establish the population connectivity among putative chromophore ionization states we employed a series of NMR saturation transfer experiments. In Fig. S6A a low power 400-ms pulse was applied to the A-state resonance at 162 ppm (red) and in another experiment at 192 ppm as an equidistant control (blue). Only the on-resonance pulse caused a loss of all B-state magnetization as evidenced by the disappearance of the peak at 177 ppm. This shows that the species at 177 ppm is interconverting with that in the broad peak at 162 ppm many times during the saturation pulse thus losing magnetic polarization. The rest of the spectrum was unaffected.

The same experiment was repeated on the other variant with slow-exchange characteristics, *s7+*:GFP S65T (Fig. S6B). The black trace is the 1D ^{13}C-NMR spectrum, and red and blue are the spectra following 400-ms saturation at the resonances at 166.4 ppm and 177.2 ppm, respectively. The fact that peaks being irradiated directly disappear entirely indicates that the pulse is sufficiently long for saturation. More importantly, the saturation also leads to depletion of the distant peak to confirm chemical exchange between them. The diminution of the downfield shoulder of the backbone carbonyl band is likely due to overlap with the saturation pulse. This interpretation is further supported by the fact that the saturation at 166.4 ppm (red trace) does not lead to a decrease in the carbonyl region. If the broad decrease between ~177-173 ppm were due instead to a broad chemical exchange feature of the chromophore one would expect to observe a decrease in magnetization in this region upon 166.4 ppm saturation, which is not seen.
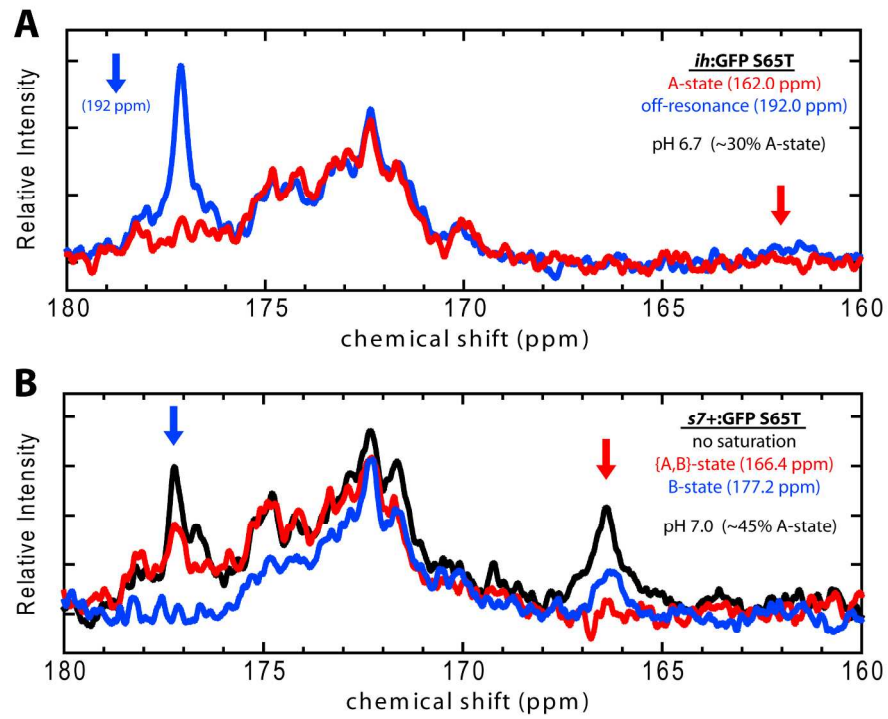
*Fig. S6*

NMR saturation transfer spectra for *ih*:GFP S65T at pH 6.7 (**A**) and *s7+*:GFP S65T at pH 7.0 (**B**).

## S.7　NMR Selective Inversion Transfer

The time-resolved selective inversion experiment observes the change in intensity of the

B-state magnetization as a function of the chemical mixing time since the A-state band has very

poor signal-to-noise as a result of exchange broadening.  If a single-site titration framework is

assumed and the $T_1$ contribution neglected then the equation for the apparent chemical exchange

is given by,

(Eq. S6)
$$M_{BZ}(t) = M_{B0}[(2p_B - 2p_B{}^2)\, e^{-t/\tau_C} + (2p_B{}^2 - p_B)]$$

where $M_{B0}$ is the equilibrium magnetization magnitude due to the B-state population, $p_B$ is the

fractional population in the B-state, and $\tau_C$ is the chemical exchange lifetime at a given pH.

Recognizing the fact that $K_a = k_{deprot}\, /\, k_{prot}$ , the chemical exchange lifetime ($\tau_C$) may be

expressed in terms of $p_B$ and $k_{deprot}$ by manipulation of Eq. S5,

(Eq. S7)
$$\tau_C(pH) = \frac{10^{pH - pK_a}}{k_{deprot}\,(1 + 10^{pH - pK_a})} = \frac{p_B(pH)}{k_{deprot}}$$

The fractional B-state population which was calculated from the UV/Vis spectrum for the pH

6.70 sample to be 79% (see SI §S.2).  In practice the triplicate data was fit to Eq. S6 to obtain $\tau_C$

which was then used to calculate $k_{deprot}$ with Eq. S7 to get 74 s$^{-1}$.

In addition to the data shown in Fig. 3D for pH 6.7, kinetic selective inversion

experiments were also conducted for *ih*:GFP S65T at pH 8.0 and 6.3.  The latter two conditions

were measured once for each mixing time while the former was done in triplicate.  Fig. S7 shows

the model from Eq. S6 applied globally.  Aside from the deprotonation constant calculated from

the single exponential fit to the pH 6.7 data in blue there are no additional free parameters.  The

values for $p_B$ were calculated from the visible absorbance spectra using the procedure in SI §S.2 and the apparent chemical exchange lifetimes were calculated as $\tau_C = p_B/k_{deprot}$ with $k_{deprot} = 74\ s^{-1}$. The relative magnitude of the B-state magnetization was determined by taking the integrated area of the peak at 177 ppm, subtracting the un-labeled blank to remove the backbone carbonyl overlap, and then normalizing to the area of the pH 8.0 peak with 0 ms mixing time.

It is notable how closely the model matches the data given the parameter constraints. The fact that the pH 8.0 sample experiences no loss of magnetization over the longest mixing time is entirely consistent with a non-exchanging deprotonated state in the limiting case of high pH. Comparing the pH 6.7 and 6.3 samples, the difference in net magnetization is slight when the mixing time is very short, reflective of the relatively small difference in the fraction deprotonated ($p_B$). However, the long time limit sees a much greater divergence in intensity preceded by differing rates of decay. Both of these features are quantitatively predicted from the global model and highlight the data's internal consistency.

The figure plots Relative Magnitude B-state Magnetization Vector versus mixing time (ms). Legend:

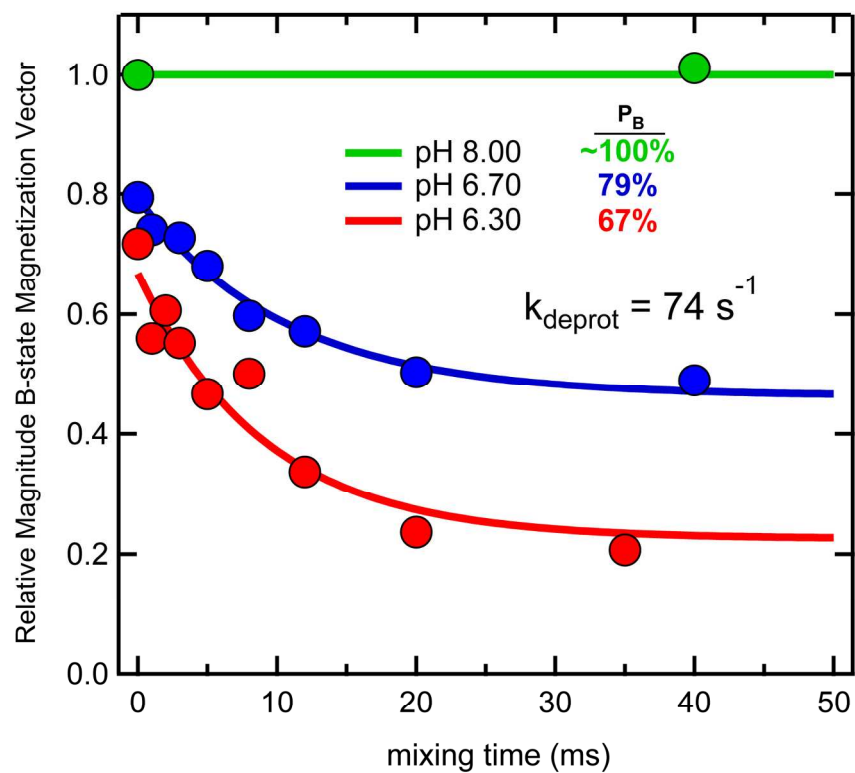| | $P_B$ |
|---|---|
| pH 8.00 | ~100% |
| pH 6.70 | 79% |
| pH 6.30 | 67% |

$k_{deprot} = 74\ s^{-1}$

*Fig. S7*

Global model of NMR kinetic selective inversion data of *ih*:GFP S65T at pH 8.0, 6.7, and 6.3 according to Eq. S6. The sole fit parameter ($k_{deprot}$) was derived from the single exponential decay constant for the pH 6.7 data (blue).

## S.8    Two-Site Kinetic Model

As introduced in the text, a two-site model of proton transfer can be expressed through a kinetic master equation.  It is possible to construct a rate matrix, $\underline{\underline{K}}$, where $K_{ij}$ is the rate constant from the state $i$ to state $j$ with the diagonal given by $K_{ii} = -\sum_{j \neq i} K_{ij}$ to ensure detailed balance.  The connectivity and rates of conversion between the four states arising from this treatment are represented by a 4x4 kinetic rate matrix ($\underline{\underline{K}}$) written explicitly below.

(Eq. S8)

$$\underline{\underline{K}} = \begin{bmatrix} -k''_{B'B}*10^{-pH} - k''_{B'A}*10^{-pH} & k''_{B'B}*10^{-pH} & k''_{B'A}*10^{-pH} & 0 \\ k'_{BB'} & -k'_{BB'} - k'_{BA} - k''_{BA'}*10^{-pH} & k'_{BA} & k''_{BA'}*10^{-pH} \\ k'_{AB'} & k'_{AB} & -k'_{AB'} - k'_{AB} - k''_{AA'}*10^{-pH} & k''_{AA'}*10^{-pH} \\ 0 & k'_{A'B} & k'_{A'A} & -k'_{A'B} - k'_{A'A} \end{bmatrix}$$

The above rate matrix $\left( \underline{\underline{K}} \right)$ is written in terms of the microscopic unimolecular rate constants for deprotonation (′) and bimolecular rate constants for protonation (″).  The parameters in Table 1, however, are given in terms of intrinsic $pK_a$'s, the energetic coupling $W$, and effective rate constants $k_{XY}$ that describe the equilibrium relaxation rate at the pH equal to the $pK_a$ for each particular microscopic step.  Specifically, the microscopic $pK_a$ for each process is shown in Fig. 6 in terms of $pK_{A'A}$, $pK_{A'B}$, and $W$.  If one considers a single acid dissociation reaction, the forward rate is given by the unimolecular deprotonation rate constant ($k'_{deprot}$) while the reverse reaction rate is the product of the bimolecular protonation rate constant and the proton concentration ($k''_{prot}[H^+] = k''_{prot}*10^{-pH}$).  $K_a$ is the ratio of the forward and reverse rate

constants ($K_a = k'_{deprot}/k''_{prot}$).  Using the definition of the $pK_a$ each of the following three

equalities can be written,

(Eq. S9a) $10^{-pK_a} = \dfrac{k'_{deprot}}{k''_{prot}}$    (Eq. S9b) $k'_{deprot} = k''_{prot} * 10^{-pK_a}$    (Eq. S9c) $k''_{prot} = k'_{deprot} * 10^{pK_a}$

It can be seen that the protonation reaction rate ($k''_{prot} * 10^{-pH}$) is equal to Eq. S9b when the pH is

equal to the $pK_a$.[*]  Furthermore, the effective relaxation rate toward equilibrium is the sum of the

forward and reverse rates.  As shown above these pseudo-unimolecular rates are the same when

the pH is equal to the $pK_a$.  Thus, the effective rate is simply twice $k'_{deprot}$, and $k''_{prot}$ is

determined by substitution into Eq. S9c.  Collectively this treatment allows us to create the rate

matrix ($\underline{\underline{K}}$) using the microscopic $pK_a$'s and effective rate constants from Table 1 while

maintaining detailed balance.

From this rate matrix, $\underline{\underline{K}}$, one may calculate the transition probability matrix as

$\underline{\underline{T}}(\tau) = \exp(\underline{\underline{K}} * \tau)$.  Each element of the transition matrix, $T_{ij}$, describes the probability that the

system is found in state $j$ after a time interval $\tau$ when starting in state $i$.  The left eigenvectors

($\vec{q}_i$) of the transition matrix and their associated eigenvalues ($\lambda_i$) provide important equilibrium

and dynamical information about the chemical network.  There is always an eigenvector having

an eigenvalue of exactly 1 which describes the relative equilibrium concentrations of all

chemical species.  The other eigenvectors report on transition modes within the network and—to

ensure mass balance—have elements that sum to zero.  Roughly speaking, the transition modes

---

[*] Note that the internal transfer between A and B is pH independent.  In this case the equilibrium constant only describes a unimolecular intramolecular proton transfer.

can describe the kinetic clustering of states of like-sign and the equilibrium chemical flux through the transition network between clusters of opposite-sign. The associated eigenvalues have values between 0 and 1 and specify the implied timescale associated with these modes given by the equation,

(Eq. S10)
$$\tau_i^* = \frac{-\tau}{\ln \lambda_i}$$

in which $\tau$ is the time interval for the transition matrix, $\underline{\underline{T}}(\tau)$, and $\lambda_i$ is the eigenvalue of the $i^{th}$ eigenvector. Specifically, modes having eigenvalues near 0 are very fast while those with eigenvalues near 1 are slow.[9]

All thermodynamic and kinetic information about the system is contained within $\underline{K}$—or equivalently $\underline{\underline{T}}(\tau)$—and can be projected onto the observables measured in FCS and NMR. Below we describe the numerical methods used for calculating the theoretical FCS autocorrelation functions and NMR spectra.

### S.8.a *FCS Autocorrelation Functions*

The equations for the autocorrelation function decay were adapted from Pramanik and Widengren.[10] Under the assumption that the diffusion coefficients of all protonation states are equal, the ACF can be expressed as the product of the diffusional and chemical exchange components, $G(\tau) = G_{dif}(\tau) * \mathrm{X}(\tau)$. The separable chemical exchange part is given by,

(Eq. S11)
$$\mathrm{X}(\tau) = \frac{\displaystyle\sum_{i=1}^{M}\sum_{j=1}^{M} Q_i Q_j q_i^{eq} T_{ij}(\tau)}{\displaystyle\sum_{i=1}^{M} Q_i q_i^{eq}}$$

where $M$ is the number of chemical species, $Q_i$ is the relative brightness of the $i^{th}$ chemical species, $T_{ij}(\tau)$ are elements of the transition probability matrix, and $q_i^{eq}$ is the equilibrium concentration of the $i^{th}$ chemical species. In the absence of any evidence to the contrary we have just assumed that both deprotonated states of the chromophore, B and B′, have the same brightness which we arbitrarily set to one and that both protonated states, A and A′, have no fluorescence. In practice, the final predicted ACF was the product of the empirically determined diffusional and dark triplet parts with the chemical exchange as in Eq. S4.

### S.8.b *NMR Spectra Predictions*

The 1D-$^{13}$CNMR spectra were numerically calculated in the time domain from the coupled Bloch-McConnell equations.[11] J-coupling does not play a significant role in the system in this study so the density matrix treatment is not necessary and this simpler approach is entirely satisfactory. In matrix form the Bloch-McConnell equation is given by,

(Eq. S12)
$$\frac{d\,\vec{M}(t)}{dt} = -\left(i\underline{\underline{L}} + \underline{\underline{R}} - \underline{\underline{K}}\right) * \vec{M}(t)$$

where $\vec{M}(t)$ is the vector of magnetizations due to each chemical species, $i$ is $\sqrt{-1}$, $\underline{\underline{L}}$ is a diagonal matrix with the frequency of each chemical species along the diagonal, $\underline{\underline{R}}$ is the relaxation matrix with the $R_2$ relaxation parameters for each species along the diagonal, and $\underline{\underline{K}}$ is the rate matrix as defined above. The solution to Eq. S12 can be written as,

(Eq. S13)
$$\vec{M}(t) = \exp\left(-\left(i\underline{\underline{L}} + \underline{\underline{R}} - \underline{\underline{K}}\right)t\right)\vec{M}(0)$$

We use the relative equilibrium concentration of all species as the initial magnetization, $\vec{M}(0) = \vec{q}^{\,eq}$. The net magnetization vector was calculated as the sum of all elements of the vector $\vec{M}(t)$. The frequency-domain signal was calculated with a fast Fourier transform of the discrete time-domain signal.

The B-state basis chemical shift used for all calculations was 177.15 ppm as determined from the peak position of the *ih*:GFP S65T sample at pH 8.0. The A-state basis was estimated from the linear extrapolation of the fit to the fraction ionized versus the peak chemical shift of the *int*PT series to be 162.00 ppm (see Fig. 3B). No distinction in chemical shift was made between A' and A or B' and B since the frequency is assumed to be dominated by the protonation state of the adjacent hydroxyl group. On account of the similar linewidths for the *ih*:GFP S65T pH 8.0 peak (~100% B-state) and the *s7*:GFP peak (~5% B-state) we assumed that the intrinsic linewidths of the A and B forms are nearly equal and used a full-width at half maximum of 0.22 ppm. Most of the NMR measurements were performed with a Varian Inova 300 MHz

spectrometer. Therefore all of the simulations presented in Fig. 7 were performed for 7.0 T at which the carbon-13 Larmor frequency is approximately 75 MHz.
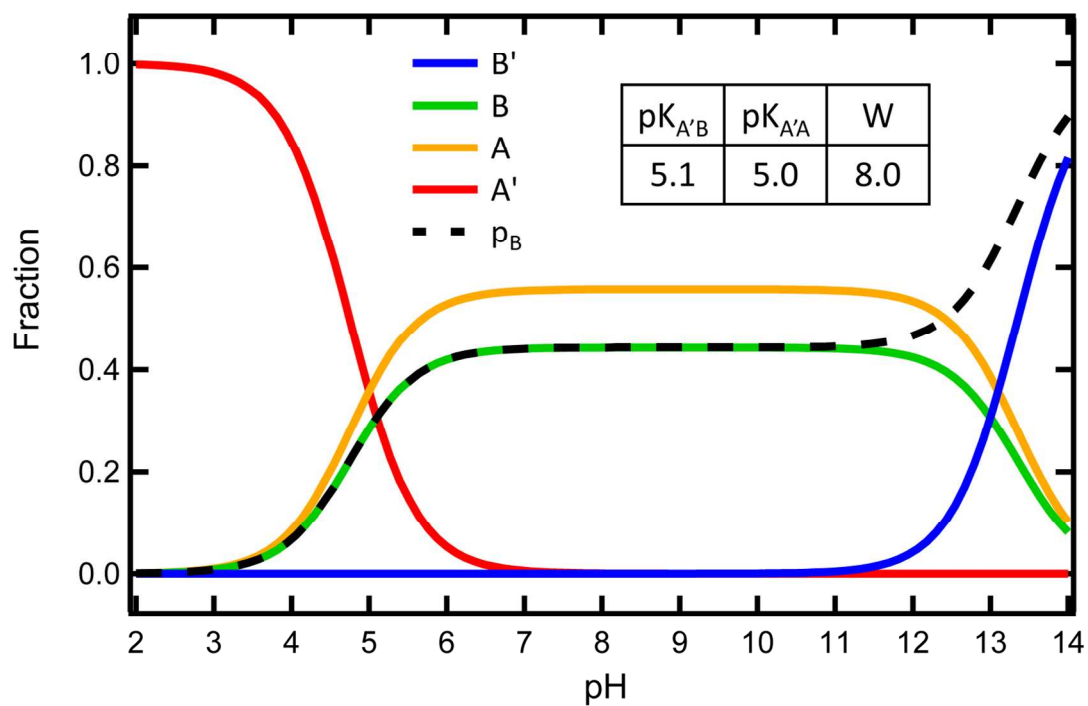
*Fig. S8*

Example titration curve for a prototypical *int*PT protein (compare to *ih*:GFP S65S in Fig. 2C). Over the range from ~6 to 12 the states B and A are dominant with their ratio determined by the difference of the intrinsic $pK_a$'s. Specifically, $\dfrac{[B]}{[A]} = 10^{-(pK_{A'A} - pK_{A'B})}$

### S.9    Titration of E222Q

Due to its established role as the terminal proton acceptor for ESPT in wtGFP[12] as well as being one of the most proximate ionizable groups to the GFP chromophore, E222 has attracted attention as a likely candidate for site "X" in Fig. 6.[13, 14] We made the isosteric mutation E222Q to assess the effect of removing the contribution of this ionizable site. Previously it was established that E222Q in wtGFP causes a breakdown in ESPT and results in a titratable *ext*PT protein superficially similar to S65T.[15, 16]

A master buffer solution was prepared with 20mM citric acid, 20mM sodium phosphate, 10mM glycine, and 150mM NaCl all adjusted to the appropriate pH by additions of 1M NaOH. Concentrated protein samples were diluted by approximately 100-fold into each buffer and the absorbance spectra were measured. The deprotonated fraction was calculated according to Eq. S3 in SI §S.2. The resulting data was fit to the single-site titration equation given in Fig. S9.

Fig. S9 shows absorbance pH titrations of *ih*:GFP S65T, *s7*:GFP S65T, and *s7+*:GFP S65T with the addition of E222Q. All of these proteins are fit reasonably well to a single-site titrations model and do not show the telltale signs of coupled sites such as negative cooperativity or non-zero A-state fractions in high pH limits as can be seen in Fig. 2C. Interestingly, while *ih*:GFP (green) and *s7+*:GFP (blue) have very similar $pK_a$'s near 5.1, *s7*:GFP (red) has a much higher $pK_a$ of 7.0. This value is getting closer to the $pK_a$ of the model chromophore HBDI of 7.8[17] and further highlights the sensitivity of the chromophore to nearby structural disturbances.
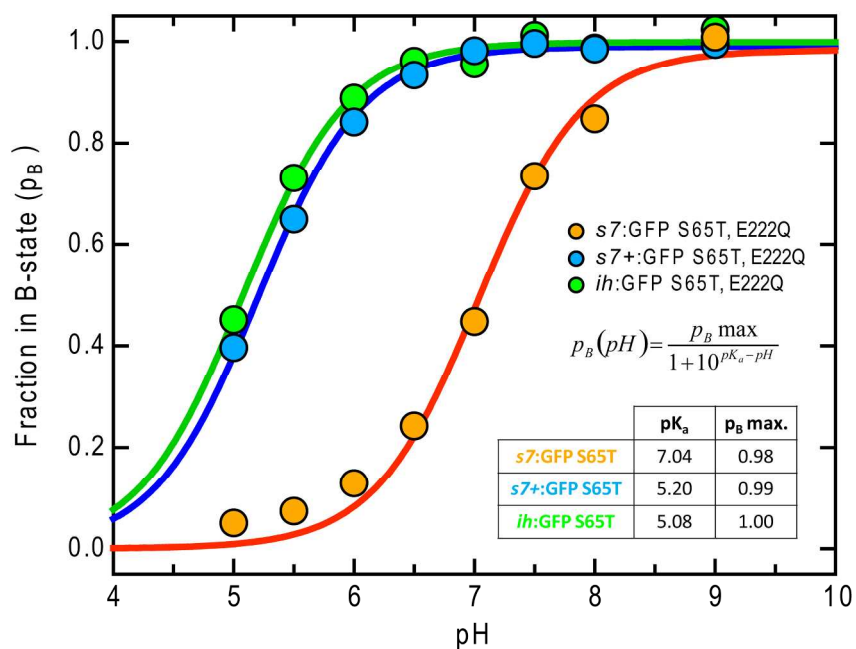
*Fig. S9*

B-state fraction calculated from the UV/vis spectra as a function of pH for selected E222Q mutants. The solid lines are fits to the single-site titration equation inset. The values for the $pK_a$'s and limiting B-state fractions ($p_B$) are given in the inset table.

### S.10    Local His-tag Effects

Hexa-histidine tags are routinely used for efficient affinity-based purification of recombinant proteins.  The effect of the His-tag is often assumed to be negligible with respect to the protein structure and function[18] and many of the counterexamples are a result of binding site occlusion.

We examined the effect of relocation and proteolytic removal of the His-tag on the two *s7* CPs (*s7*:GFP and *s7+*:GFP) due to its relative proximity to the chromophore.  We found that the presence of the His-tag on the N-terminus gave rise to a significant shift in the apparent chromophore $pK_a$ towards higher values.  In the case of *s7*:GFP S65T and *s7+*:GFP S65T the magnitude of these shifts were 0.5 and 0.8 pH units, respectively (see Fig. S10A).

In our view, there are two general modes in which His-tags might affect the $pK_a$ of a buried ionizable group.  Firstly, the positive charge borne by the His-tag could electrostatically favor chromophore deprotonation.  Secondly, the His-tag may act as a reservoir of protons gathered from solution thus causing an apparent increase in their local concentration.  Our observation that presence of the His-tag favors the protonated chromophore form inclines us towards this latter explanation.  The electrostatic effects are likely of less consequence due to the solvent exposure of the His-tag and the accompanying charge shielding by water and counterions.  If we posit a simple local proton enrichment factor $\alpha$, such that $[H^+]_{local} = \alpha[H^+]_{sol}$, we can write an expression for the modified $pK_{a\ obs}$,

(Eq. S14) $$pK_{a\ obs} = pK_{a\ in} + \log\alpha$$

$pK_{a\ obs}$ is the observed $pK_a$ in the presence of the His-tag, $pK_{a\ in}$ is the intrinsic $pK_a$ in the absence of the His-tag, and $\alpha$ is the local proton enrichment factor which will, in general, be a function of pH.

This simple analysis can be further extended to estimate the consequences for chemical exchange kinetics if the same microscopic rate constants are retained and the factor $\alpha$ is treated as a constant over this approximately one unit pH window. It is straightforward to show that the apparent rate of chemical exchange at a pH equal to the $pK_a$ for an ionization step to be,

(Eq. S15) $$k_{app} = 2\,k_{deprot}$$

If the factor $\alpha$ is included for a His-tag perturbation, algebraic manipulation leads to an expression for the apparent chemical exchange at a pH equal to the new perturbed $pK_a$ (see Eq. S14) in the presence of the His-tag,

(Eq. S16) $$k_{app} = k_{deprot}\left(1 + \frac{1}{\alpha}\right)$$

In the limit of no local pH perturbation (i.e. $\alpha = 1$) the result is identical to Eq. S15. However, with $\alpha > 1$ the prediction is that chemical exchange rate in the presence of the perturbation is slower than in its absence.

The predicted slowing effect is evident in our data and can be seen most clearly by comparison of the B-state (~177ppm) behavior upon pH titration (see Fig. S10 B&C). Specifically, in $s7+$:GFP S65T with the close His-tag, the peak retains its width and position and merely decreases in intensity with decreasing pH. In contrast the peak in $s7+$:GFP S65T lacking

the His-tag decreases in intensity but simultaneously increases dramatically in width and begins
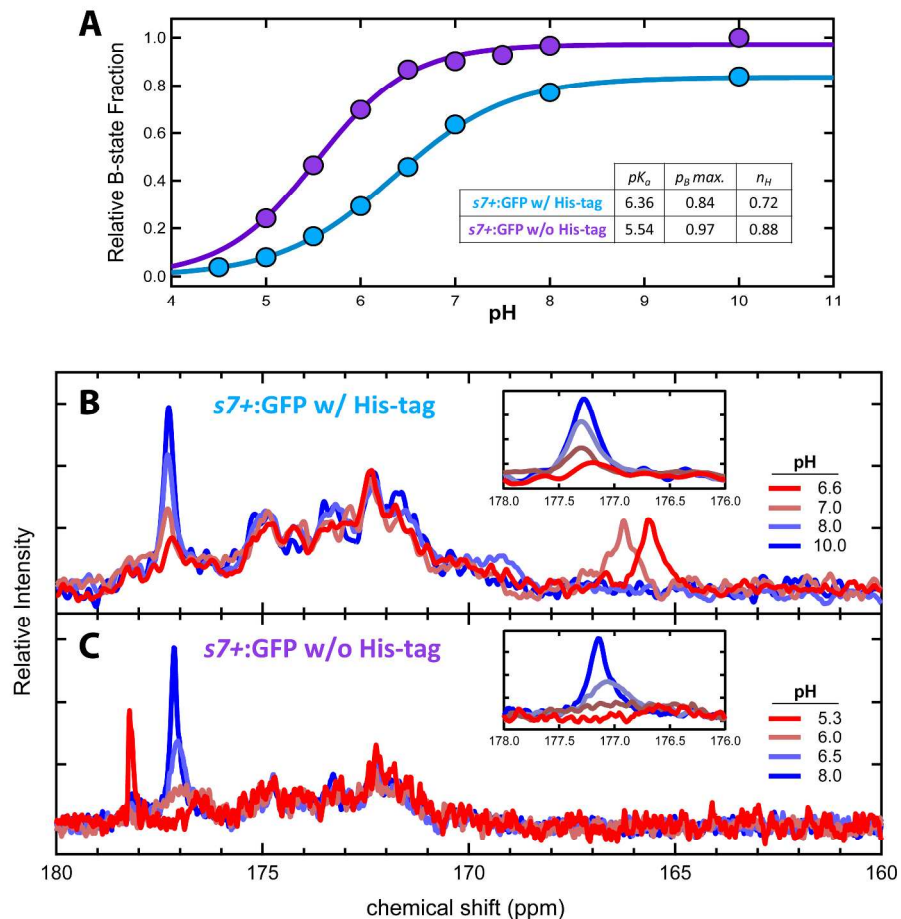
to shift upfield.

*Fig. S10*

**A)** pH titrations of *s7+*:GFP S65T with and without an N-terminal hexa-histidine affinity tag. The solid lines are fits to a Hill titration equation of the form, $p_B(pH) = \dfrac{p_B \max}{1 + 10^{n_H(pK_a - pH)}}$ where $p_B$ *max* is the fraction in the B-state in the high pH limit, the $pK_a$ is the pH value at which the $p_B$ is half the value of $p_B$ *max*, and $n_H$ is the Hill coefficient describing the extent of cooperativity (values less than one are negatively cooperative). The fit parameters are shown in the inset table. **B)** 1D $^{13}$C-NMR spectra of *s7+*:GFP S65T with the His-tag collected at 2.0 mM on a 300 MHz NMR. **C)** 1D $^{13}$C-NMR spectra of *s7+*:GFP S65T without the His-tag collected at 0.3 mM on a 500 MHz NMR (the peak near 178 ppm at pH 5.3 is due to the citrate buffer).

**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, *Nature Biotechnology 24*, 79-88.

(2) Ward, W. W., Prentice, H. J., Roth, A. F., Cody, C. W., and Reeves, S. C. (1982) Spectral Perturbations of the Aequorea Green-Fluorescent Protein, *Photochemistry and Photobiology 35*, 803-808.

(3) Richarz, R., and Wuthrich, K. (1978) C-13 Nmr Chemical-Shifts of Common Amino-Acid Residues Measured in Aqueous-Solutions of Linear Tetrapeptides H-Gly-Gly-X-L-Ala-Oh, *Biopolymers 17*, 2133-2141.

(4) Grimsley, G. R., Scholtz, J. M., and Pace, C. N. (2009) A summary of the measured pK values of the ionizable groups in folded proteins, *Protein Science 18*, 247-251.

(5) Ulrich, E. L., Akutsu, H., Doreleijers, J. F., Harano, Y., Ioannidis, Y. E., Lin, J., Livny, M., Mading, S., Maziuk, D., Miller, Z., Nakatani, E., Schulte, C. F., Tolmie, D. E., Kent Wenger, R., Yao, H., and Markley, J. L. (2008) BioMagResBank, *Nucleic Acids Research 36*, D402-D408.

(6) Phillips, G. N. (1997) Structure and dynamics of green fluorescent protein, *Current Opinion in Structural Biology 7*, 821-827.

(7) Zacharias, D. A., Violin, J. D., Newton, A. C., and Tsien, R. Y. (2002) Partitioning of lipid-modified monomeric GFPs into membrane microdomains of live cells, *Science 296*, 913-916.

(8)  Mizuno, H., Mal, T. K., Walchli, M., Fukano, T., Ikura, M., and Miyawaki, A. (2010) Molecular basis of photochromism of a fluorescent protein revealed by direct (13)C detection under laser illumination, *Journal of Biomolecular Nmr 48*, 237-246.

(9)  Noe, F., and Fischer, S. (2008) Transition networks for modeling the kinetics of conformational change in macromolecules, *Current Opinion in Structural Biology 18*, 154-162.

(10)  Pramanik, A., and Widengren, J. (2006) Fluorescence Correlation Spectroscopy (FCS), In *Encyclopedia of Molecular Cell Biology and Molecular Medicine*, Wiley-VCH Verlag GmbH & Co. KGaA.

(11)  McConnell, H. M. (1958) Reaction rates by nuclear magnetic resonance, *Journal of Chemical Physics 28*, 430-431.

(12)  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.

(13)  Bizzarri, R., Nifosi, R., Abbruzzetti, S., Rocchia, W., Guidi, S., Arosio, D., Garau, G., Campanini, B., Grandi, E., Ricci, F., Viappiani, C., and Beltram, F. (2007) Green fluorescent protein ground states: the influence of a second protonation site near the chromophore, *Biochemistry 46*, 5494-5504.

(14)  Scharnagl, C., Raupp-Kossmann, R., and Fischer, S. F. (1999) Molecular basis for pH sensitivity and proton transfer in green fluorescent protein: protonation and conformational substates from electrostatic calculations, *Biophys J 77*, 1839-1857.

(15) 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.

(16) Kent, K. P., Childs, W., and Boxer, S. G. (2008) Deconstructing green fluorescent protein, *Journal of the American Chemical Society 130*, 9664-+.

(17) Conyard, J., Kondo, M., Heisler, I. A., Jones, G., Baldridge, A., Tolbert, L. M., Solntsev, K. M., and Meech, S. R. (2011) Chemically Modulating the Photophysics of the GFP Chromophore, *Journal of Physical Chemistry B 115*, 1571-1576.

(18) Carson, M., Johnson, D. H., McDonald, H., Brouillette, C., and DeLucas, L. J. (2007) His-tag impact on structure, *Acta Crystallographica Section D-Biological Crystallography 63*, 295-301.