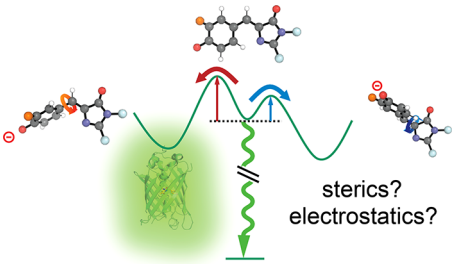
Energetic Basis and Design of Enzyme Function Demonstrated Using GFP, an Excited-State Enzyme

Chi-Yun Lin,[§] Matthew G. Romei,[§] Irimpan I. Mathews, and Steven G. Boxer*

ABSTRACT: The past decades have witnessed an explosion of *de novo* protein designs with a remarkable range of scaffolds. It remains challenging, however, to design catalytic functions that are competitive with naturally occurring counterparts as well as biomimetic or nonbiological catalysts. Although directed evolution often offers efficient solutions, the fitness landscape remains opaque. Green fluorescent protein (GFP), which has revolutionized biological imaging and assays, is one of the most redesigned proteins. While not an enzyme in the conventional sense, GFPs feature competing excited-state decay pathways with the same steric and electrostatic origins as conventional ground-state catalysts, and they exert exquisite control over multiple reaction outcomes through the same principles. Thus, GFP is an "excited-state enzyme". Herein we show that rationally designed mutants and hybrids that contain environmental mutations and substituted chromophores provide the basis for a quantitative model and prediction that describes the influence of sterics and electrostatics on excited-state catalysis of GFPs. As both perturbations can selectively bias photoisomerization pathways, GFPs with fluorescence quantum yields (FQYs) and photoswitching characteristics tailored for specific applications could be predicted and then demonstrated. The underlying energetic landscape, readily accessible via spectroscopy for GFPs, offers an important missing link in the design of protein function that is generalizable to catalyst design.



1. INTRODUCTION

Numerous methods have been employed in developing GFPs with desired behaviors,[1−17] including directed evolution and high-throughput screening of mutant libraries[5−9] that optimize brightness. Machine learning has afforded redder and brighter GFPs,[10,11] and *de novo* protein design has reduced the size of GFP.[12] Unfortunately, the former lacks physical insight, and the latter does not factor in structure−FQY relationships, leading to a FQY (∼2%) substantially below those of GFPs derived from *Aequorea victoria* (avGFP; FQY ∼ 80%). Only through further substantial screening and chromophore modification were brighter versions (FQY ∼ 23%) obtained.[13] Photoswitching, the ability to toggle between strongly and weakly fluorescent states through irradiation,[18,19] is another useful function that facilitates super-resolution imaging and optogenetic applications.[20,21] One of the most common photoswitching mechanisms is photoisomerization (Figure 1A), an excited-state bond-rotation pathway that competes with fluorescence emission. Due to this competition, selecting for an efficient photoswitchable protein is difficult via high-throughput screens; past efforts have relied on naturally occurring photoisomerizable GFPs as starting points[14] and/or painstaking combinations of rational design and screening.[15−17] A physical framework capturing the protein environmental factors that control the FQY and photoisomerization in GFPs is necessary to guide more efficient designs, and this is intimately related to the challenge of catalyst design.

In earlier work, we discovered that the FQY of the *anionic* GFP chromophore embedded in the fixed native protein environments of Dronpa2 or superfolder GFP can be modulated through the introduction of electron-donating and -withdrawing substituents.[22] Because electron-donating and -withdrawing groups red-shift and blue-shift the chromophore, respectively (Figure 2A), we can use the corresponding transition energies (derived from absorption maxima, blue vertical arrow in Figure 1A) to gauge the extents of electronic perturbation conferred by substituents, representing the changes in electrostatic interaction between the modified chromophore and the fixed environment. The FQY exhibits a peaked trend when correlated with transition energy (Figure 2A; now converted into driving force, *vide infra*), allowing us to unambiguously identify the electrostatic influences on FQYs. Since the FQYs are mainly modulated via photoisomerization rate constants (k_P and k_I in Figure 1A) rather than the spontaneous emission rate constant (k_{fl} in Figure 1A), the nonmonotonic trend reveals *two* competing nonradiative photoisomerization pathways (Figure 1A) associated with

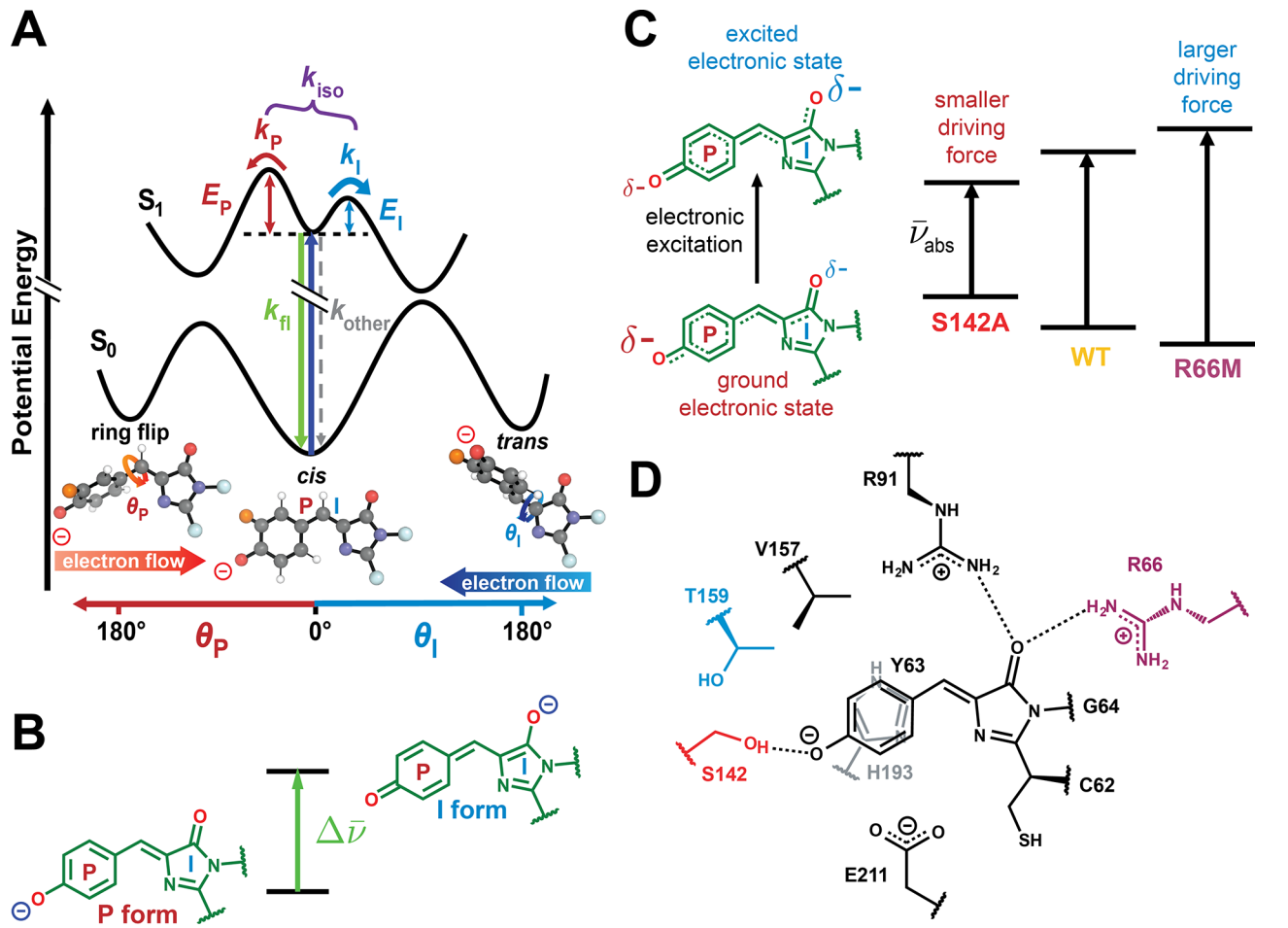


Figure 1. Energetics and local environment of the GFP chromophore in Dronpa2. (A) Potential energy surfaces (PESs) for the anionic GFP chromophore along the isomerization coordinate, modified from Figure 4C in ref 22. Copyright (2020) American Association for the Advancement of Science. After excitation from the cis ground state (indigo arrow), the chromophore can either fluoresce (k_{fl}) or decay by isomerization through excited-state barrier crossing (k_{iso}) and conical intersections (trajectory not shown) or by other nonradiative pathways (k_{other}) back to the ground state. Isomerization can either occur about the phenolate bond (P bond; k_P, phenolate ring flip) or the imidazolinone bond (I bond; k_I, cis−trans isomerization), with opposite directions of electron flow. The relative barrier heights (E_P and E_I) depend on steric and electrostatic factors of the environment around the chromophore,[22] catalyzing one pathway over the other. Note that the PESs are drawn as if there is no Stokes shift, but the nuclear displacement between the ground and excited states before and right after excitation, respectively, is dominantly along the bond-length alternation coordinate,[23] which is orthogonal to the isomerization coordinates shown here. Therefore, a nonzero Stokes shift is still present, but not displayed along the reaction coordinate as shown. (B) The driving force of the chromophore $\Delta\bar{\nu}$ is defined as the relative energy between the P (left) and the I (right) resonance forms in a given environment. In all proteins studied in this work, the P form is consistently lower in energy,[23] defined as a positive driving force. (C) Marcus−Hush model explaining shifts in transition energy $\bar{\nu}_{abs}$ depending on the electrostatic influence of the protein environment on the chromophore's ground and excited states.[23] (D) The chromophore and its local environment within Dronpa2. R66, S142, and T159 are the residues mutated in this work, while tyrosine analogues in place of Y63 are used to introduce substituents into the phenolate ring of the chromophore.[22]

opposite electron flow directions between the rings. The probability of the substituted chromophore adopting either photoisomerization pathway is influenced by the *electrostatic* interaction between the protein environment and the electron flow within the chromophore during photoisomerization,[24,25] such that one pathway is preferred over the other in the presence of electron-donating or -withdrawing groups. The opposite electron flow directions, shown as red and blue horizontal thick arrows below the ball-and-stick models in Figure 1A, can be understood by the disruption in the π-conjugated system of the chromophore caused by the twisting about the two exocyclic bonds (the P and I bonds) in the excited state, forcing electrons to redistribute within the chromophore along the isomerization coordinate.[24,25] Therefore, by interacting with the electron redistribution along the isomerization pathways, *electrostatics* can cause bond-selective photoisomerization of the chromophore, complementing the more commonly argued role of steric hindrance in suppressing chromophore (photo)isomerization.[3,26,27] The relative barrier heights E_P and E_I determine the outcome (Figure 1A), and control of these barrier heights is analogous to conventional concepts in catalysis.

While the transition energy is already a good metric for estimating the electrostatic perturbation that is free of steric influences, as the absorption process only involves electronic redistribution rather than nuclear rearrangements (a Franck−Condon process), we define the driving force $\Delta\bar{\nu}$ (Figure 1B),[23,28] which is the relative energy between the P and I resonance forms of the chromophore, to rigorously quantify electrostatics. We argue that $\Delta\bar{\nu}$ responds linearly to

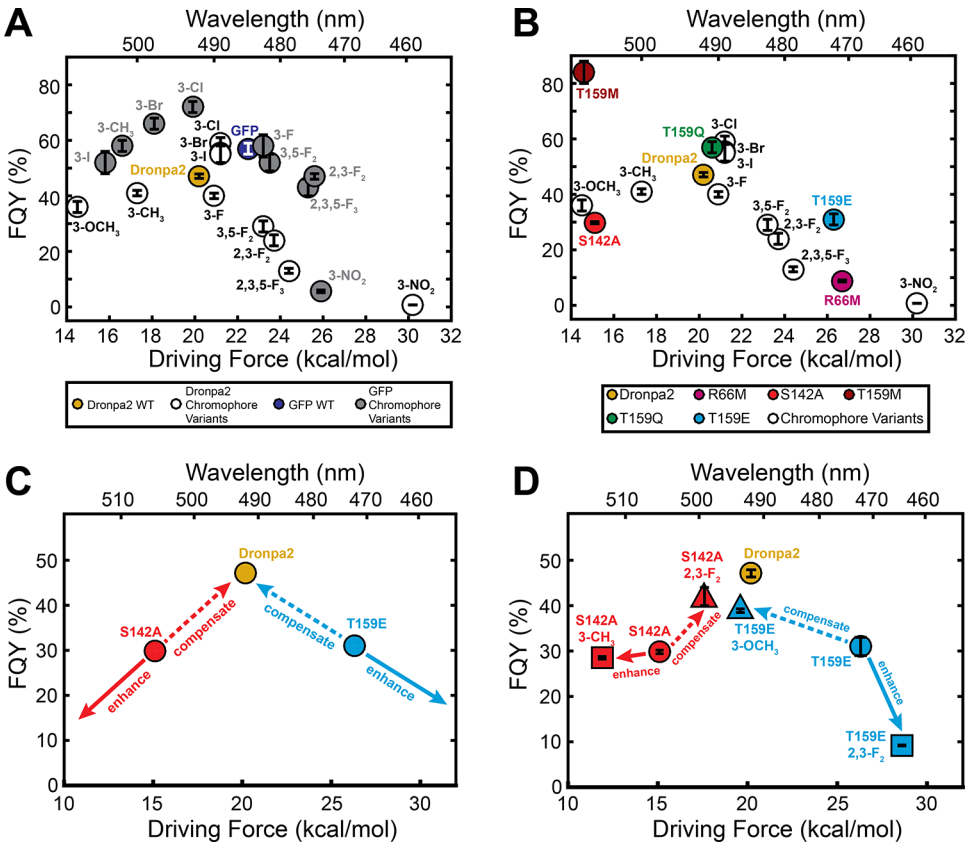


Figure 2. Correlation plots of FQY and room-temperature driving force. (A) Relationship between FQY and driving force (Figure 1C; converted from eq 1) for unsubstituted and substituted chromophores within Dronpa2 and GFP (Table S3). In both Dronpa2 and GFP, varying the electronic properties of the chromophore using substituents leads to a nonmonotonic peaked trend. (B) The dependence of FQY on the chromophore's driving force for environmental mutants (colored circles, Table S1) and chromophore variants (white) of Dronpa2 (Table S3). (C,D) The dependence of FQY on the chromophore's driving force in the Dronpa2 compensating and enhancing hybrids schematically (C) and experimentally (D), plotted based on Table 2.

electrostatic perturbations and exhibits additivity when multiple electrostatic sources are present, while the transition energy does not share these properties,[23] so the former is a preferred metric. The use of $\Delta\bar{\nu}$ is motivated by the color-tuning behavior of the anionic GFP chromophore in electrostatic fields, which can be explained by resonance color theory[28] or the more advanced Marcus–Hush treatment.[23] In these models, the electronic distribution of the anionic GFP chromophore in the ground or excited states before and right after excitation (Figure 1C, left), respectively, is described as the superposition of the P and I resonance forms (or charge-localized forms, Figure 1B) that are orthogonal to each other. Note that the ground and excited states are the energy eigenstates, while the resonance forms are not, so an electronic coupling V_0 between the two resonance forms is necessary. A nonlinear correspondence between the transition energy $\bar{\nu}_{abs}$ (i.e., the energy difference between excited and ground states) and $\Delta\bar{\nu}$ (i.e., the energy difference between the two resonance forms) can therefore be derived:[23,28]

$$\bar{\nu}_{abs} = \sqrt{(\Delta\bar{\nu})^2 + (2V_0)^2} \quad (1)$$

where V_0 is determined to be 9530 cm^{-1}, based on the correlation plots of various photophysical properties of anionic GFP chromophores covalently modified by substituents and/or embedded in a wide range of environments.[23] The value of V_0, an intrinsic property of the chromophore, is in fact not only applicable to the anionic GFP chromophore, but also shared by the photoactive yellow protein chromophore and cyanine dyes, so long as there are exactly 8 conjugated bonds separating the charge localization centers (e.g., the P- and I-ring oxygens in the anionic GFP chromophore) for a charged chromophore.[23] The qualitative meaning of eq 1 regarding the color-tuning behavior of the GFP chromophore can be described as follows: with respect to the wild-type environment or chromophore, any decrease or increase in $\Delta\bar{\nu}$ caused by modifications results in a red or blue shift, respectively (Figure 1C). The driving force can be perturbed through either direct modification of the chromophore or through changes in the protein environment, so it can serve as an ideal quantity to reflect the electron distribution of the chromophore,[23] unify both sources of perturbations,[29] and connect to the underlying theme of electrostatic catalysis. For example, in the case of modified chromophores, placing an electron-withdrawing group at the P ring stabilizes the P form more than the I form (Figure 1B) and increases the driving force relative to the unsubstituted chromophore. As the ground state inherits more of P-form character while the excited-state exhibits a stronger I-form character (Figure 1C), the transition energy becomes larger and leads to a blue-shifted absorption maximum (eq 1 and Figure 2A).

2. RESULTS AND DISCUSSION

2.1. Tuning Electrostatics with Mutants and Hybrids.

Figure 1D shows the chromophore environment of Dronpa2,

Table 1. Driving Force $\Delta\bar{\nu}$ Predictions for Each Dronpa2 Hybrid[a]

chromophore variant	driving force $\Delta\bar{\nu}$ (cm^{-1})	difference from Dronpa2 $\Delta\Delta\bar{\nu}$ (cm^{-1})	hybrid protein	point mutant driving force $\Delta\Delta\bar{\nu}$ (cm^{-1})	substituent driving force $\Delta\Delta\bar{\nu}$ (cm^{-1})	predicted combined driving force $\Delta\bar{\nu}$ (cm^{-1})
Dronpa2 ("wild type")	7010	0	S142A/ 2,3-F_2	5300	+1290 (compensating)	6590
2,3-F_2	8300	+1290	S142A/ 3-CH_3	5300	−820 (enhancing)	4480
3-CH_3	6190	−820	T159E/ 2,3-F_2	9200	+1290 (enhancing)	10490
3-OCH_3	5070	−1940	T159E/ 3-OCH_3	9200	−1940 (compensating)	7260

[a]The left side shows either the additive or subtractive effect of a particular chromophore substituent on the driving force. The right side shows the predicted driving force for each hybrid combining the effect of the point mutant and the chromophore substituent. Driving force values are extracted from ref 23. and calculated from eq 1 with an electronic coupling V_0 of 9530 cm^{-1}. The chromophore modified with OCH_3 possesses a somewhat smaller V_0 than the unsubstituted counterpart,[23] but for the current purpose the same V_0 is used for driving force evaluation.

Table 2. Predicted and Observed Driving Forces, Absorption Peak Maxima, and FQYs for Each Dronpa2 Hybrid[a]

			observed driving force $\Delta\bar{\nu}$		observed absorption peak maximum (transition energy $\bar{\nu}_{abs}$)				
hybrid protein	predicted combined driving force $\Delta\bar{\nu}$ (cm^{-1})	predicted absorption peak maximum (nm)	(cm^{-1})	(kcal/mol)	(cm^{-1})	(nm)	(kcal/mol)	FQY (%)	FQY SD (%)
T159E/2,3-F_2 (enhancing)	10490	459.6	9990	28.6	21520	464.7	61.5	9.2	0.1
T159E	N/A	N/A	9190	26.3	21160	472.5	60.5	31	2
T159E/3-OCH_3 (compensating)	7260	490.3	6870	19.6	20260	493.6	57.9	38.9	0.4
Dronpa2 ("wild type")	N/A	N/A	7070	20.2	20310	492.4	58.1	46	2
S142A/2,3-F_2 (compensating)	6590	495.9	6160	17.6	20030	499.3	57.3	42	2
S142A	N/A	N/A	5290	15.1	19780	505.5	56.6	29.8	0.4
S142A/3-CH_3 (enhancing)	4480	510.7	4170	11.9	19510	512.6	55.8	28.5	0.3

[a]The observed driving force is calculated from eq 1 with an electronic coupling V_0 of 9530 cm^{-1}.[23] SD: standard deviation. The predicted absorption peak maxima are at most 5 nm from the observed ones, with better accuracy for redder species.

which exhibits a balance between emission and photoisomerization. To isolate the electrostatic effects, residues immediately surrounding the chromophore were replaced with amino acids that minimized differences in size. The S142A mutation causes a red shift by destabilizing the P form through removal of a hydrogen bond to the phenolate oxygen (Figures 1B, 1C, S1A, and S2A). The blue-shifted R66M mutant results from I-form destabilization via the removal of the favorable electrostatic interaction between the arginine and the imidazolinone oxygen (Figures 1B, 1C, S1A, and S2B). Within an isosteric T159 mutant series (T159M, T159Q, T159E), T159M is the most red-shifted (by 15 nm compared to wild type), while increasing polarity and/or charge causes a blue shift in T159Q/E; the glutamine and glutamate in T159Q and T159E mutants, respectively, replace S142 as the primary hydrogen bonding partner to the phenolate oxygen and preferentially stabilize the P form (Figures S1A and S2C–S2F).

We next measured the FQYs (Table S1) and plotted them against the corresponding room-temperature driving forces (eq 1) to determine the electrostatic effect on photoisomerization (Figure 2A), as the room-temperature absorption maxima are readily accessible (Figure S1). S142A and R66M have a decreased FQY along with strong red- and blue-shifted peak maxima, respectively (red and purple circles in Figure 2B), recapitulating the peaked trend for chromophore variants (white circles in Figure 2B). In contrast, the isosteric T159 mutant series displays a linear correlation with peak maximum (blue, green, and maroon circles in Figure 2B), rendering Dronpa (T159M) an apparent outlier of the trend. We attribute this to a consistently increased steric effect for the isosteric series relative to wild-type Dronpa2 in conjunction with the electrostatic mechanism (Section 2.3). Nevertheless, we still find that the FQY can be tuned electrostatically through environmental mutations.

To circumvent the confounding steric effect, we created hybrids by introducing substituted chromophores into environmental mutants. We first chose one red-shifted (S142A) and one blue-shifted (T159E) mutant with the wild-type Dronpa2 chromophore. We then introduced electron-donating or -withdrawing chromophore substituents to the P ring, which would be predicted to either respectively enhance or compensate for the electronic effect of the mutant with respect to wild-type properties. For example, as the S142A mutation destabilizes the P form, an "enhancing" chromophore modification would be electron-donating and push the electronic properties of the chromophore (driving force and FQY) even further from wild type. A "compensating" modification with an electron-withdrawing group would stabilize the P form, countering the mutational effect and creating a more wild-type-like chromophore (Figure 2C). Note that the same substituent can act as enhancing or compensating in different environmental contexts according to electrostatic FQY tuning.

For the hybrids, we can quantitatively predict the optimal substituent, within the range available,[22] to pair with a given

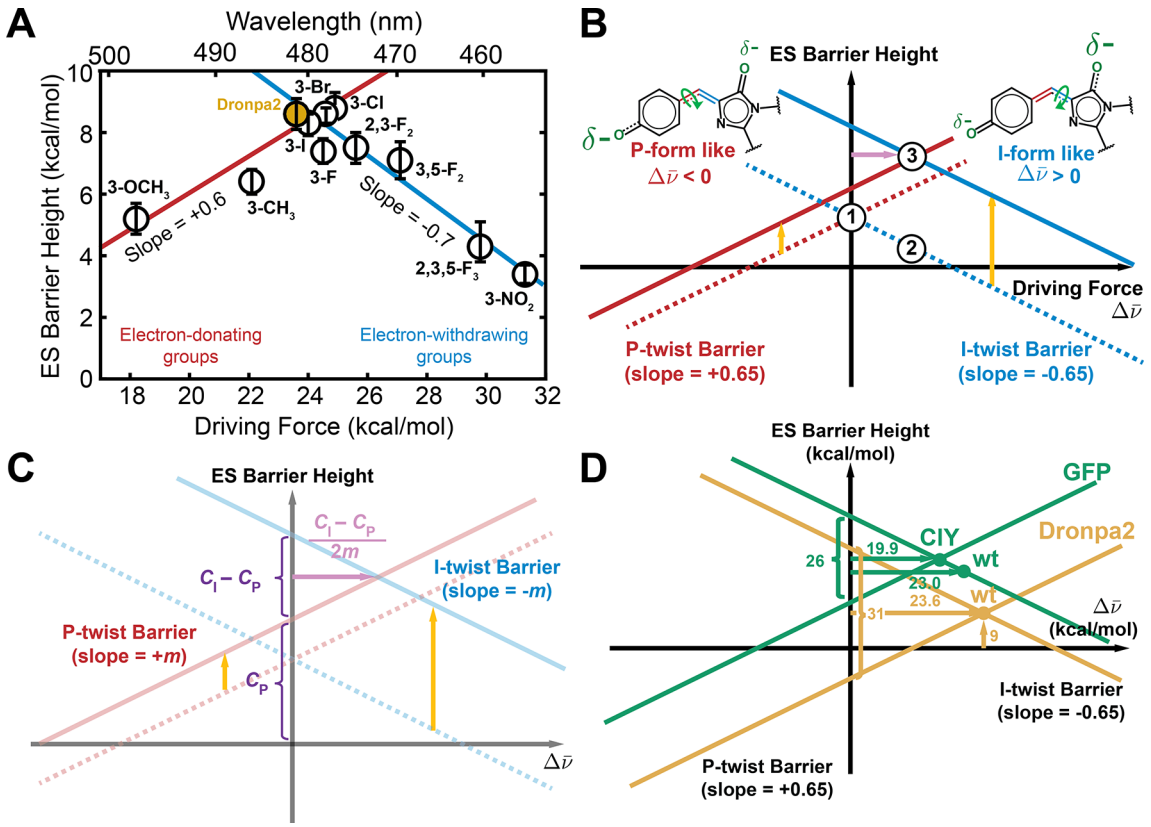


Figure 3. Qualitative and quantitative analysis of sterics and electrostatics of the GFP chromophore within a protein environment. The vertical axes, namely excited-state (ES) barrier heights, refers to *E* in eq 3. (A) Excited-state energy barriers for Dronpa2 chromophore variants plotted against 77-K driving forces (Table S3), modified from Figure 3B in ref 22. Copyright (2020) American Association for the Advancement of Science. The fit through the electron-withdrawing and -donating group points is shown as a blue and red line, respectively, with wild-type Dronpa2 shown in gold at the apex. (B) Schematic showing effects of sterics around the P and I rings of the chromophore on the magnitude and apex position of the excited-state energy barrier, shown as blue and red lines for I and P twist, respectively. Without steric effects (dashed lines), the apex lies at zero driving force (case 1). Shifting the driving force to positive energy (i.e., to the right) leads to a preference for I-bond rotation due to a lower barrier (case 2). Greater steric confinement (solid lines) around the I ring (i.e., longer yellow arrow for I-twist than P-twist) causes the apex to shift to the right (positive driving forces, case 3). (C) The algebraic relationship between the apex shift and differential sterics according to panel B. (D) Interplay between steric and electrostatic effects for the excited-state barrier height of GFP (green) and Dronpa2 (gold). The driving forces are inferred from 77-K absorption maxima (Table S3). See also Figure 5A.

mutant based on driving force additivity (Table 1). Each point mutant has a driving force, to which a fixed value is added or subtracted based on the chromophore substituent, obtained from the difference between the driving force of Dronpa2 with a natural and substituted chromophore.[23] For the compensating hybrids, the optimal substituents to bring the driving force of S142A and T159E close to wild type are 2,3-F_2 and 3-OCH_3, respectively. For the enhancing hybrids, we chose substituents with low steric bulk but that still provide a large perturbation to the driving force: S142A/3-CH_3 and T159E/2,3-F_2. The observed absorption peak maximum for each hybrid agrees well with the predictions (Table 2; Figures S1B and S1C): incorporation of electron-donating and -withdrawing substituents leads to the predicted red and blue shift, respectively. Figure 2D shows the correlation between FQY and driving force for the Dronpa2 hybrids. Both enhancing hybrids (S142A/3-CH_3 and T159E/2,3-F_2) have a decreased FQY, pushing the values further from wild type as anticipated from electrostatic FQY tuning (squares in Figure 2D). Remarkably, both compensating hybrids (S142A/2,3-F_2 and T159E/3-OCH_3) have an increased FQY compared to the respective mutant with the unsubstituted chromophore, bringing the values closer to the wild-type value (triangles in Figure 2D). This observation implies that the electronic effect of the chromophore substituent successfully compensates for the electrostatic perturbation caused by the environmental mutation. Either the chromophore substituents (2,3-F_2 or 3-OCH_3) or the environmental mutations (S142A or T159E) alone each cause a decrease in FQY compared to the wild-type Dronpa2, so the observation of an increased FQY in these compensating hybrids suggests cooperativity ("reciprocal sign epistasis")[8,30] between deleterious perturbations that cannot otherwise be explained without electrostatic FQY tuning.

2.2. Predictive Model for Steric and Electrostatic Effects on Excited-State Catalysis. The FQY φ_{fl} is the ratio between the intrinsic spontaneous emission rate k_{fl} and the total excited-state decay rate constants[31] (Figure 1A):

$$\varphi_{fl} = k_{fl}\tau = \frac{k_{fl}}{k_{fl} + k_{iso} + k_{other}} \quad (2)$$

where k_{iso} and k_{other} denote the total rate constant for excited-state isomerization and other nonradiative pathways, respectively; τ is the fluorescence lifetime. We can then dissect the temperature, electrostatic, and steric dependence of each term to understand how the chromophore's FQY is influenced by its environment. k_{fl} is minimally tunable through electrostatics as

evidenced by the nearly constant transition dipole moment across different GFP mutants;[23,32] steric effects are irrelevant since emission involves electronic redistribution but barely any nuclear rearrangement (another Franck−Condon process). The only way the protein environment can tune the FQY is through modulating the competing nonradiative decay pathways. k_{fl} is estimated to be $(3.5\ \text{ns})^{-1}$,[33] which should be applicable to the unsubstituted GFP chromophore in any environment, so any nonradiative process much slower than this value cannot tune FQYs.

k_{other} arises from both direct internal conversion and intersystem crossing, but the latter is much less competitive than other excited-state processes[34] (see also Section S2 regarding the substituted chromophores). Accordingly, we can approximate k_{other} with a single rate constant from direct internal conversion k_{IC} due to vibrational wave function overlap between the ground and excited electronic states, which is relatively temperature insensitive (see Section S6 of ref 22 and Section S11 of ref 31). To obtain k_{IC}, we examine a GFP mutant series in which the threonine at position 203 is replaced with aromatic side chains that $\pi-\pi$ stack with the chromophore P ring and can be varied in electron richness. The corresponding FQYs are nearly constant around 77% despite the modified electrostatic interaction (Figure S3 and Table S2). Steric hindrance by the aromatic ring overwhelms electrostatics and renders k_{iso} uncompetitive; the remaining 23% of excited-state decay can be ascribed to internal conversion; k_{IC} is $(12\ \text{ns})^{-1}$ and imposes an upper limit for GFP's FQY of approximately 80%,[35] close to that of avGFP. Extensive mutational studies also demonstrate that avGFP is indeed located at the local maximum of the fitness landscape for brightness.[8] Any approach that slows excited-state isomerization down to tens of nanoseconds is sufficient to maximize FQY.

In contrast with other processes, excited-state isomerization requires crossing over an energy barrier along with significant electronic and nuclear motion (Figure 1A), so the isomerization rate k_{iso} is almost solely responsible for the temperature, electrostatic, and steric dependence of FQY.[22] The associated barriers are typically >3 kcal/mol for GFPs,[22] and the corresponding rate constants are comparable with k_{fl} (ns time scale). The rapid intramolecular vibrational energy redistribution (ps time scale)[31,39] right after excitation renders the system thermally equilibrated before emission and isomerization, so the assumption for Arrhenius behavior, also common for ground-state catalysis, is met for isomerization. A pre-exponential factor A and an energy barrier E can thus be assigned for each isomerization pathway:

$$k_{iso} = A_P \exp\left(-\frac{E_P}{RT}\right) + A_I \exp\left(-\frac{E_I}{RT}\right) \approx A \exp\left(-\frac{E}{RT}\right) \tag{3}$$

where k_{iso} is then approximated with a single Arrhenius expression when we measure the excited-state energy barrier E of Dronpa2 variants using the temperature dependence of their fluorescence lifetimes.[22] A_P and A_I are close in value as experimentally verified,[22,40] so A should be close to both A_P and A_I, and the measured excited-state barrier height E can well approximate the lesser of the two barriers, E_P or E_I (Figure 1A). A is 10^3 to 10^5 ns^{-1},[22] agreeing well with the value estimated from transition state theory ($kT/h \sim 10^{13}\ \text{s}^{-1}$). This suggests that when the excited-state barrier exceeds 9 kcal/mol (i.e., k_{iso} being 1% of k_{fl} at 300 K), as for the $\pi-\pi$ stacking GFP mutants (Figure S3), no further increase in FQY can be seen as it reaches the upper limit.

We now replot the excited-state barriers from Dronpa2 variants (Figure 3B in ref 22) against the corresponding driving forces to better understand the electrostatic effect (Figure 3A). Note that from now on, we will estimate the driving forces based on their 77-K absorption maxima, as we believe that they are preferable estimates for electrostatic energetic contributions and the corresponding values are better resolved across halogenated Dronpa2s than the room-temperature counterparts (Table S3). Linear fits to the electron-donating and -withdrawing substituent data exhibit slopes of +0.6 and −0.7, reflecting the electrostatic sensitivity of E_P and E_I, respectively. The assignment of the P- and I-bond rotation to each value will be explicated later. These slopes are about equal in magnitude (∼ 0.65 within experimental errors) and opposite in sign; the signs agree well with a model treating the chromophore as an allylic anion, in which the two rings of the anionic GFP chromophore were approximated as two p orbitals (Figure 4 in ref 22). Analogous to electrostatic enzyme catalysis,[41,42] this electrostatic sensitivity originates from chromophore charge redistribution during photoisomerization interacting with the protein environment (Figure 1A), effectively an excited-state enzyme that selectively catalyzes either P- or I-bond rotation. We expect these slopes in Figure 3A to be directly transferable to different environments around the chromophore, since the driving force is the only parameter responsible for the electrostatic sensitivity of the entire PES:[22]

$$E_P = 0.65\Delta\bar{\nu} + C_P \quad \text{and} \quad E_I = -0.65\Delta\bar{\nu} + C_I \tag{4}$$

The linear approximation is evidently a simplification; however, it is sufficient for the observed driving force range (Figure 3A) and its simplicity can already afford insights, as for linear free-energy relationships (i.e., linear dependence of free energy barriers on net free energies) in physical organic chemistry[43] and the Butler−Volmer equation (i.e., linear dependence of free energy barriers on applied potentials) in electrochemistry.[44] Steric effects, including the intrinsic barrier to bond isomerization in the absence of any external steric constraint, can be separated out in terms of empirical constants C_P and C_I. We can then rewrite eqs 2 and 3 to explicitly show the electrostatic and steric dependence of the FQY:

$$\varphi_{fl}(T, \Delta\bar{\nu}, \text{sterics}) \approx \frac{k_{fl}}{k_{fl} + k_{other} + A\left[\exp\left(-0.65\frac{\Delta\bar{\nu}}{RT} - \frac{C_P(\text{sterics})}{RT}\right) + \exp\left(0.65\frac{\Delta\bar{\nu}}{RT} - \frac{C_I(\text{sterics})}{RT}\right)\right]} \tag{5}$$

Two factors mediate excited-state pathway selection: sterics, which acts upon large scale nuclear motion of two rings during isomerization, and electrostatics, which interacts with electronic redistribution during isomerization (or driving force). The electrostatic influence of the red fluorescent protein environment on the corresponding chromophore's FQY is also extensively discussed by a recent paper,[45] while our physical model treats electrostatics differently and

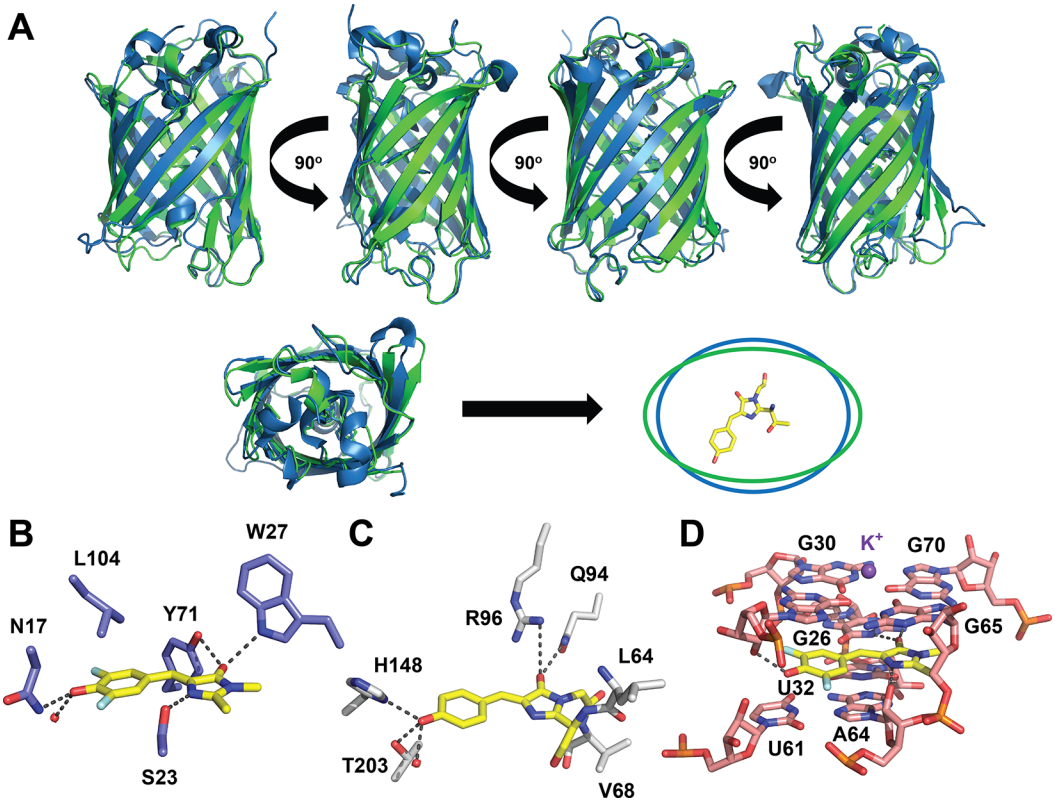


Figure 4. GFP chromophore (yellow) in various biomolecular environments. (A) Overlaid β-barrels of Dronpa2 (green, PDB: 6NQJ) and GFP (blue, PDB: 6OFK). The barrels are shown in different perspectives to illustrate the differences in dimensions. The overlaid ovals at the right bottom corner, color coded according to the proteins they represent, are exaggerated simplification for the cross sections of the barrels. P-twist motion clashes with residues along the wider dimension, for which GFP is tighter than Dronpa2. (B) mFAP1 (PDB: 6CZI). (C) avGFP (PDB: 2WUR). (D) Spinach (PDB: 4TS2). In panels B–D, hydrogen bonds associated with the chromophore are shown as dashed lines.

explicitly incorporates the steric component (see Section S3 in Supporting Information). A mixture of electrostatic and steric effects has also been proposed and observed in other photobiological systems, such as rhodopsins.[46] According to eq 5, FQY is a nonlinear function of $\Delta\bar{\nu}$, and thus the linear additivity of driving force does not translate to an additivity of FQY, as observed from the compensating hybrids (Figure 2D and Table 2). Cooperativity between mutations, a phenomenon that renders protein design and even directed evolution challenging,[30,47] could similarly be partly explained by a nonlinear function (i.e., FQY) encoding two (or more) pathways dependent on an additive underlying parameter (i.e., driving force).[23] Steric effects C_P and C_I serve as an alternative tuning mechanism for the excited-state barriers E_P and E_I, preventing the FQY from being completely tied to color via electrostatics, as is the case for other photophysical properties.[23] If C_P equals C_I, there should be no preference for either isomerization pathway when $\Delta\bar{\nu} = 0$, corresponding to a maximum FQY (eq 5; Figure 3B, case 1). Since $\Delta\bar{\nu} = 0$ also corresponds to the reddest possible absorption (eq 1), a combination of these two equations would suggest that the redder the chromophore, the higher the FQY by varying $\Delta\bar{\nu}$. However, we observe an apex in the trend that is not centered at $\Delta\bar{\nu} = 0$ (Figure 3A), suggesting that C_P is not identical with C_I. Intuitively, the volume-demanding I twist experiences more steric hindrance than the P twist within the protein environment since the I ring is covalently anchored.

With eq 4, we can explain the apex position in the FQY (or excited-state barrier) vs driving force plot (Figure 3B). The sign of the driving force is defined positive when the P form is more stable than the I form, which is the case for all proteins studied so far[23] (Figure 1B). With zero differential sterics from the protein environment ($C_P = C_I$; dashed lines) and zero driving force, the negative charge of the anionic chromophore is maximally delocalized and both exocyclic bonds are equally probable to twist upon excitation. This corresponds to the largest possible barrier when $C_P = C_I$, and the apex is located at $\Delta\bar{\nu} = 0$ (Figure 3B, case 1). When the driving force becomes positive (right side of Figure 3B), which means the P form is lower in energy than the I form, electron density is reduced at the I bond (i.e., more single-bond character; more I-form like) upon excitation (Figure 1C), and the I twist becomes more favorable (Figure 3B, case 2), as in the case of the chromophore in vacuum, whose driving force is positive due to the larger intrinsic electron affinity of P ring compared to that of I ring.[48] The correlation between the driving force and the isomerization tendency in the excited state was also shown via modeling the system as an allylic anion.[22,49] If the I ring is anchored inside the protein, C_I becomes larger than C_P (yellow vertical arrows and solid lines in Figure 3B). Consequently, the apex shifts along the x-axis and lies at a positive driving force, as observed in Figure 3A, and it also increases along the y-axis due to the resulting constriction on bond rotation (Figure 3B, case 3). At that apex, the driving force from electrostatic influences matches the apex shift caused by differential steric interactions. However, when the steric effects are large enough to render k_{iso} uncompetitive with k_{fl} (Figure S3), the maximally allowed FQY is reached, and the apex for FQY cannot be

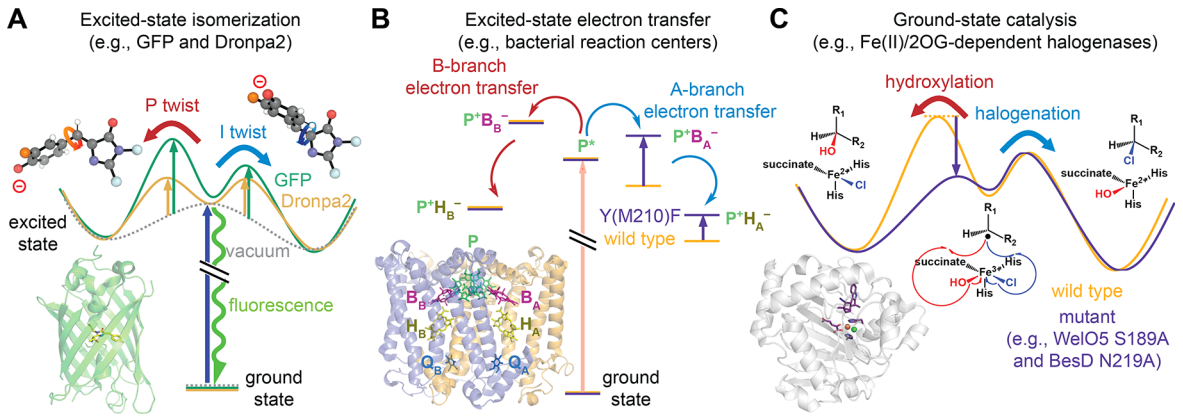


Figure 5. Energetic control of competing pathways for excited- and ground-state catalysis by diverse protein environments. (A) GFP (green) and Dronpa2 (gold) protein environments suppress excited-state isomerization of the chromophore to different degrees compared to that in vacuum (gray), rendering GFP less photoisomerizable than Dronpa2 (Figure 3D). (B) Y(M210)F mutant (purple) of *Rhodobacter sphaeroides* photosynthetic reaction center reveals that tyrosine at M210, which stabilizes the first intermediate, is in part responsible for the unidirectional excited-state electron transfer of wild type (orange).[62,63] (C) Wild-type Fe(II)/2-oxoglutarate (2OG)-dependent halogenases (orange) chlorinate their substrates, but their intrinsic hydroxylating power can be unleashed upon mutation (purple).[60,64,65] The default (blue) and the side pathways (red for all and green for panel A) are shown on the right and left for each panel, respectively. Energies are not drawn to scale.

detected. Note that the driving force at the apex is determined from the *differential* sterics ($C_I - C_P$), while the barrier heights are affected by the *absolute* sterics (C_I or C_P), so it is possible to have an apex location at zero driving force when steric hindrance to the P twist is comparable with I ring anchoring (Figure 3C).

2.3. Applications, Generalizations, and Implications for Design. This model allows us to quantitatively evaluate the contributions of sterics and electrostatics to excited-state catalysis. In particular, it informs us how Dronpa2 is superior in photoisomerization compared to GFP dissected in terms of sterics and electrostatics, leading to the former being a better candidate for superresolution microscopy. From Figure 3A, the correlation plot of excited-state barriers and driving forces, wild-type Dronpa2 sits at the apex among all Dronpa2 variants. As its FQY (~ 50%; gold circle in Figure 2A) is far from the maximally allowed 80%, this implies that the corresponding driving force (23.6 kcal/mol; Figure 3A) offsets the differential sterics, so we can estimate the differential sterics as 31 kcal/mol (23.6 × 2 × 0.65, Figure 3C and gold vertical bracket in Figure 3D). Even though we do not have a corresponding plot for superfolder GFP as in Figure 3A, the apex (the monochlorinated variant, gray "3-Cl" in Figure 2A) of the FQY vs driving force plot lies at a driving force of 19.9 kcal/mol at 77 K (Table S3) and approaches the FQY limit of 80%.[22] The corresponding differential sterics, shared by wild-type Dronpa2 and its monochlorinated variant, is 26 kcal/mol (= 19.9 × 2 × 0.65, Figure 3C and green vertical bracket in Figure 3D). Combined with the fact that GFP has a higher apex FQY than Dronpa2 (blue and gold circles in Figure 2A), we can infer that the overall steric contribution should be higher for GFP than Dronpa2, but the differential sterics is also 5 kcal/mol smaller (= 31−26) for GFP (gold and green vertical brackets Figure 3D), leading to an apex located at a smaller driving force than Dronpa2 (Figure 3D). This is explained by a tighter β-barrel for GFP compared to Dronpa2, resulting in a more sterically hindered P twist (Figure 4A). Moreover, since the unmodified chromophore in the GFP environment possesses a driving force of 23.0 kcal/mol (Table S3, as opposed to 23.6 kcal/mol for that in Dronpa2), the Dronpa2 I-twist barrier is also lowered by 0.4 kcal/mol (= (23.6−23.0) × 0.65) electrostatically compared to the GFP counterpart. Therefore, both steric and electrostatic (to a lesser extent, Section S4) effects work together in the GFP barrel to promote chromophore fluorescence, while Dronpa2 exhibits a higher photoisomerization efficiency (Figure 5A). While it would be desirable to extract the *absolute* contributions of electrostatics and sterics from the protein environments with respect to the chromophore in vacuum, we argue that this is not possible with the available data without further unvalidated assumptions (see Section S4). For the Dronpa2 T159 isosteric series, the consistently lengthened side chain for M, Q, or E within the series compared to that of T in wild-type Dronpa2 creates more steric bulk to P twist and shifts the apex to a smaller driving force and higher FQY (Figure 2B, see also Figure S2 for X-ray structures). This is why the T159 isosteric mutants exhibit higher FQYs compared to other Dronpa2 mutants or variants with similar driving forces and why T159M appears as a significant outlier to the peaked trend (Figure 2B).

This analysis can also explain why the *de novo* designed mFAPs (Figure 4B) failed to recapitulate avGFP's high FQYs (Figure 4C)[12] and more generally how an understanding of the energy landscape can provide guidance for the design of functional proteins. Original mFAPs utilize the same difluorinated chromophore as the RNA mimic Spinach (Figure 4D)[50] to encourage chromophore deprotonation, but fluorines lower the I-twist barrier as electron-withdrawing substituents (Figure 3A).[22] In Spinach, $\pi-\pi$ stacking with G-quadruplexes effectively inhibits isomerization (Figure 4D),[50,51] leading to a FQY of 72%. In mFAPs, however, the chromophore is neither anchored to the protein as in avGFP (Figure 4C) nor motionally restricted. M27W is present in mFAP1 and mFAP2 to interact with the I ring via a hydrogen bond (Figure 4B), but this interaction is not sufficient to restore the maximal FQY. To further increase the FQYs, this analysis suggests the addition or removal of fluorines from the chromophore's I or P rings, respectively, and the introduction of aromatic amino acids near the chromophore's P ring to encourage $\pi-\pi$ stacking interactions. In fact, the newly installed $-CF_3$ group on the I ring and L104H likely explains the much-improved FQY (23%) of chromophore-bound mFAP10.[13]

3. CONCLUSIONS

GFP is both green and fluorescent, while the free GFP chromophore in water is neither, so it is tempting to ascribe this drastic change in properties solely to the protein environment. However, the chromophore's ability to be green and fluorescent is already encoded in its PESs (i.e., energy landscape), but is unveiled by protein environments through creating energetic barriers for excited-state isomerization, and these properties can also be elicited using nonprotein environments.[3,27] An analogous example is the relationship between an enzyme and its substrate. The availability of different reaction pathways and the potential for pathway selection, existing for numerous ground-state and excited-state enzymes,[52−55] are primarily inscribed in the PES(s) of the chromophore/substrate and constrain how the PES(s) can be tuned in response to an environment,[56] illustrated by diverse examples in Figure 5. On the basis of directed evolution studies on enzymes, new chemistries are not created out of nowhere but rather the substrates are found to already exhibit low reactivities toward the said chemis-tries.[57−60] Therefore, to rationally design enzymes that are superior at catalyzing a reaction, it is important to sample a wide range of perturbations to substrates (or chromophores capable of structural change) and the environment's steric or electrostatic influences on the energetics of nonproductive yet competitive pathways rather than only those that exhibit more desirable phenotypes.[61] Only when those less desirable cases are understood can we mechanistically deduce why the more productive pathway is not taken, guiding future design efforts to optimize the desired function.

To uncover the role of the protein environment on GFP photoisomerization, our approach of dissecting the underlying energetics is partially enabled by noncanonical amino acids as perturbations, which either subtly interpolate between effects from canonical amino acids or extrapolate to uncharted territories. Functional extrapolation has been largely emphasized with recent advancements in genetic codon expansion,[66] while interpolation has been rather overlooked in comparison. It is the ability to prepare subtle mutants or variants that facilitates deeper mechanistic studies and energetic probing, in combination with rigorous characterizations, such as spectroscopic, kinetic, functional, and structural analyses and theoretical modeling. Subtle electrostatic/electronic and steric modulation to critical structural elements and/or physical properties for protein functions, including but not limited to cation−π interactions,[67] π−π interactions,[23] hydrogen bonds,[68,69] dipole−π interactions,[63] pK_a,[70] nucleophilicity,[71] electronic distribution,[22] redox potential,[72] active-site electric field,[73] and isomerization efficiency,[74] can then be quantified by a suitable parameter that is derived from the physical chemistry or physical organic[43] tradition. The explored sequence space can accordingly be encoded with such parameter to interrogate its importance on energetics during catalysis, providing insights into the design principles of the large classes of enzymes adopting polar or radical mechanisms.

AUTHOR INFORMATION

Corresponding Author

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

Authors

Chi-Yun Lin − *Department of Chemistry, Stanford University, Stanford, California 94305, United States;* Present Address: Department of Chemistry, The Pennsylvania State University, University Park, Pennsylvania 16802, United States; orcid.org/0000-0001-6555-8767

Matthew G. Romei − *Department of Chemistry, Stanford University, Stanford, California 94305, United States;* Present Address: Antibody Engineering, Genentech Inc., South San Francisco, California 94080, United States.; orcid.org/0000-0001-6798-7493

Irimpan I. Mathews − *Stanford Synchrotron Radiation Lightsource, Menlo Park, California 94025, United States;* orcid.org/0000-0001-6254-3519

Author Contributions

[§]C.-Y.L. and M.G.R. contributed equally to this work.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank Jacob M. Kirsh and Steven D. E. Fried for useful comments. C.-Y.L. was supported by a Kenneth and Nina Tai Stanford Graduate Fellowship and the Taiwanese Ministry of Education. M.G.R. was supported by a Center for Molecular Analysis and Design graduate fellowship. This work was supported, in part, by NIH Grant GM118044 (to S.G.B.) and NSF CCI Phase I: Center for First-Principles Design of Quantum Processes (CHE-1740645). Use of the Stanford Synchrotron Radiation Lightsource (SSRL), SLAC National Accelerator Laboratory, is supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract No. DE-AC02-76SF00515. The SSRL Structural Molecular Biology Program is supported by the DOE Office of Biological and Environmental Research, and by the National Institutes of Health, National Institute of General Medical Sciences (including P41GM103393). The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official views of NIGMS or NIH. Part of this work was performed at the Stanford Nano Shared Facilities (SNSF), supported by the National Science Foundation under award ECCS-1542152.

REFERENCES

(1) Chudakov, D. M.; Matz, M. V.; Lukyanov, S.; Lukyanov, K. A. Fluorescent proteins and their applications in imaging living cells and tissues. *Physiol. Rev.* **2010**, *90*, 1103−1163.

(2) Nienhaus, K.; Ulrich Nienhaus, G. Fluorescent proteins for live-cell imaging with super-resolution. *Chem. Soc. Rev.* **2014**, *43*, 1088−1106.
(3) Walker, C. L.; Lukyanov, K. A.; Yampolsky, I. V.; Mishin, A. S.; Bommarius, A. S.; Duraj-Thatte, A. M.; Azizi, B.; Tolbert, L. M.; Solntsev, K. M. Fluorescence imaging using synthetic GFP chromophores. *Curr. Opin. Chem. Biol.* **2015**, *27*, 64−74.
(4) Ferré-D'Amaré, A. R.; Truong, L. From fluorescent proteins to fluorogenic RNAs: Tools for imaging cellular macromolecules. *Protein Sci.* **2019**, *28*, 1374−1386.
(5) Shaner, N. C.; Campbell, R. E.; Steinbach, P. A.; Giepmans, B. N. G.; Palmer, A. E.; Tsien, R. Y. Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma sp.* red fluorescent protein. *Nat. Biotechnol.* **2004**, *22*, 1567−1572.
(6) Pédelacq, J.-D.; Cabantous, S.; Tran, T.; Terwilliger, T. C.; Waldo, G. S. Engineering and characterization of a superfolder green fluorescent protein. *Nat. Biotechnol.* **2006**, *24*, 79−88.
(7) Yoo, T. H.; Link, A. J.; Tirrell, D. A. Evolution of a fluorinated green fluorescent protein. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 13887−13890.
(8) Sarkisyan, K. S.; Bolotin, D. A.; Meer, M. V.; Usmanova, D. R.; Mishin, A. S.; Sharonov, G. V.; Ivankov, D. N.; Bozhanova, N. G.; Baranov, M. S.; Soylemez, O.; Bogatyreva, N. S.; Vlasov, P. K.; Egorov, E. S.; Logacheva, M. D.; Kondrashov, A. S.; Chudakov, D. M.; Putintseva, E. V.; Mamedov, I. Z.; Tawfik, D. S.; Lukyanov, K. A.; Kondrashov, F. A. Local fitness landscape of the green fluorescent protein. *Nature* **2016**, *533*, 397−401.
(9) Manna, P.; Hung, S.-T.; Mukherjee, S.; Friis, P.; Simpson, D. M.; Lo, M. N.; Palmer, A. E.; Jimenez, R. Directed evolution of excited state lifetime and brightness in FusionRed using a microfluidic sorter. *Integr. Biol.* **2018**, *10*, 516−526.
(10) Saito, Y.; Oikawa, M.; Nakazawa, H.; Niide, T.; Kameda, T.; Tsuda, K.; Umetsu, M. Machine-learning-guided mutagenesis for directed evolution of fluorescent proteins. *ACS Synth. Biol.* **2018**, *7*, 2014−2022.
(11) Alley, E. C.; Khimulya, G.; Biswas, S.; AlQuraishi, M.; Church, G. M. Unified rational protein engineering with sequence-based deep representation learning. *Nat. Methods* **2019**, *16*, 1315−1322.
(12) Dou, J.; Vorobieva, A. A.; Sheffler, W.; Doyle, L. A.; Park, H.; Bick, M. J.; Mao, B.; Foight, G. W.; Lee, M. Y.; Gagnon, L. A.; Carter, L.; Sankaran, B.; Ovchinnikov, S.; Marcos, E.; Huang, P.-S.; Vaughan, J. C.; Stoddard, B. L.; Baker, D. De novo design of a fluorescence-activating β-barrel. *Nature* **2018**, *561*, 485−491.
(13) Klima, J. C.; Doyle, L. A.; Lee, J. D.; Rappleye, M.; Gagnon, L. A.; Lee, M. Y.; Barros, E. P.; Vorobieva, A. A.; Dou, J.; Bremner, S.; Quon, J. S.; Chow, C. M.; Carter, L.; Mack, D. L.; Amaro, R. E.; Vaughan, J. C.; Berndt, A.; Stoddard, B. L.; Baker, D. Incorporation of sensing modalities into de novo designed fluorescence-activating proteins. *Nat. Commun.* **2021**, *12*, 856.
(14) Ando, R.; Mizuno, H.; Miyawaki, A. Regulated fast nucleocytoplasmic shuttling observed by reversible protein high-lighting. *Science* **2004**, *306*, 1370−1373.
(15) Stiel, A. C.; Andresen, M.; Bock, H.; Hilbert, M.; Schilde, J.; Schönle, A.; Eggeling, C.; Egner, A.; Hell, S. W.; Jakobs, S. Generation of monomeric reversibly switchable red fluorescent proteins for far-field fluorescence nanoscopy. *Biophys. J.* **2008**, *95*, 2989−2997.
(16) Grotjohann, T.; Testa, I.; Leutenegger, M.; Bock, H.; Urban, N. T.; Lavoie-Cardinal, F.; Willig, K. I.; Eggeling, C.; Jakobs, S.; Hell, S. W. Diffraction-unlimited all-optical imaging and writing with a photochromic GFP. *Nature* **2011**, *478*, 204−208.
(17) El Khatib, M.; Martins, A.; Bourgeois, D.; Colletier, J.-P.; Adam, V. Rational design of ultrastable and reversibly photo-switchable fluorescent proteins for super-resolution imaging of the bacterial periplasm. *Sci. Rep.* **2016**, *6*, 18459.
(18) Acharya, A.; Bogdanov, A. M.; Grigorenko, B. L.; Bravaya, K. B.; Nemukhin, A. V.; Lukyanov, K. A.; Krylov, A. I. Photoinduced chemistry in fluorescent proteins: Curse of blessing? *Chem. Rev.* **2017**, *117*, 758−795.
(19) Bourgeois, D.; Adam, V. Reversible photoswitching in fluorescent proteins: A mechanistic view. *IUBMB Life* **2012**, *64*, 482−491.
(20) Shcherbakova, D. M.; Sengupta, P.; Lippincott-Schwartz, J.; Verkhusha, V. V. Photocontrollable fluorescent proteins for super-resolution imaging. *Annu. Rev. Biophys.* **2014**, *43*, 303−329.
(21) Zhou, X. X.; Lin, M. Z. Photoswitchable fluorescent proteins: Ten years of colorful chemistry and exciting applications. *Curr. Opin. Chem. Biol.* **2013**, *17*, 682−690.
(22) Romei, M. G.; Lin, C.-Y.; Mathews, I. I.; Boxer, S. G. Electrostatic control of photoisomerization pathways in proteins. *Science* **2020**, *367*, 76−79.
(23) Lin, C.-Y.; Romei, M. G.; Oltrogge, L. M.; Mathews, I. I.; Boxer, S. G. Unified model for photophysical and electro-optical properties of green fluorescent proteins. *J. Am. Chem. Soc.* **2019**, *141*, 15250−15265.
(24) Olsen, S.; Lamothe, K.; Martínez, T. J. Protonic gating of excited-state twisting and charge localization in GFP chromophores: A mechanistic hypothesis for reversible photoswitching. *J. Am. Chem. Soc.* **2010**, *132*, 1192−1193.
(25) Martin, M. E.; Negri, F.; Olivucci, M. Origin, nature, and fate of the fluorescent state of the green fluorescent protein chromophore at the CASPT2//CASSCF resolution. *J. Am. Chem. Soc.* **2004**, *126*, 5452−5464.
(26) Stiel, A. C.; Trowitzsch, S.; Weber, G.; Andresen, M.; Eggeling, C.; Hell, S. W.; Jakobs, S.; Wahl, M. C. 1.8 Å bright-state structure of the reversibly switchable fluorescent protein Dronpa guides the generation of fast switching variants. *Biochem. J.* **2007**, *402*, 35−42.
(27) Tolbert, L. M.; Baldridge, A.; Kowalik, J.; Solntsev, K. M. Collapse and recovery of green fluorescent protein chromophore emission through topological effects. *Acc. Chem. Res.* **2012**, *45*, 171−181.
(28) Olsen, S. A modified resonance-theoretic framework for structure−property relationships in a halochromic oxonol dye. *J. Chem. Theory Comput.* **2010**, *6*, 1089−1103.
(29) Marder, S. R.; Gorman, C. B.; Meyers, F.; Perry, J. W.; Bourhill, G.; Brédas, J.-L.; Pierce, B. M. A unified description of linear and nonlinear polarization in organic polymethine dyes. *Science* **1994**, *265*, 632−635.
(30) Poelwijk, F. J.; Kiviet, D. J.; Weinreich, D. M.; Tans, S. J. Empirical fitness landscapes reveal accessible evolutionary paths. *Nature* **2007**, *445*, 383−386.
(31) Lin, C.-Y.; Both, J.; Do, K.; Boxer, S. G. Mechanism and bottlenecks in strand photodissociation of split green fluorescent proteins (GFPs). *Proc. Natl. Acad. Sci. U.S.A.* **2017**, *114*, E2146−E2155.
(32) However, when introducing substituents to the chromophore, the transition dipole moment and the extinction coefficient change due to covalent modification. Changes in these parameters likely result in the observed asymmetric FQY trend (Figure 2A) when plotted against transition energy for Dronpa2 chromophore variants, while the corresponding excited-state barrier trend is symmetric (Figure 3A). Specifically, the methoxy modified chromophore (far red-shifted) has a higher k_{fl} than the nitro and trifluoro modified chromophores (far blue-shifted).
(33) Striker, G.; Subramaniam, V.; Seidel, C. A. M.; Volkmer, A. Photochromicity and fluorescence lifetimes of green florescent protein. *J. Phys. Chem. B* **1999**, *103*, 8612−8617.
(34) Byrdin, M.; Duan, C.; Bourgeois, D.; Brettel, K. A long-lived triplet state is the entrance gateway to oxidative photochemistry in green fluorescent proteins. *J. Am. Chem. Soc.* **2018**, *140*, 2897−2905.
(35) One caveat of this statement is that the normal modes and their corresponding vibronic couplings are different for the chromophore in a vacuum vs protein,[36] so k_{IC} in each case could be drastically different. Here we assume that since low-frequency vibrational modes are more important than higher frequency ones for internal conversion due to larger numbers of participating phonons,[37,38] k_{IC} should be roughly constant across protein environments with β-barrel scaffolds.

(36) Bochenkova, A. V.; Andersen, L. H. Ultrafast dual photoresponse of isolated biological chromophores: link to the photoinduced mode-specific non-adiabatic dynamics in proteins. *Faraday Discuss.* **2013**, *163*, 297−319.
(37) Englman, R.; Jortner, J. The energy gap law for radiationless transitions in large molecules. *Mol. Phys.* **1970**, *18*, 145−164.
(38) Nitzan, A. *Chemical Dynamics in Condensed Phases*, 1st ed.; Oxford University Press: New York, 2006; pp 439−449.
(39) Oscar, B. G.; Zhu, L.; Wolfendeen, H.; Rozanov, N. D.; Chang, A.; Stout, K. T.; Sandwisch, J. W.; Porter, J. J.; Mehl, R. A.; Fang, C. Dissecting optical response and molecular structure of fluorescent proteins with non-canonical chromophores. *Front. Mol. Biosci.* **2020**, *7*, 131.
(40) Drobizhev, M.; Hughes, T. E.; Stepanenko, Y.; Wnuk, P.; O'Donnell, K.; Scott, J. N.; Callis, P. R.; Mikhaylov, A.; Dokken, L.; Rebane, A. Primary role of the chromophore bond length alternation in reversible photoconversion of red fluorescence proteins. *Sci. Rep.* **2012**, *2*, 688.
(41) Fried, S. D.; Bagchi, S.; Boxer, S. G. Extreme electric fields power catalysis in the active site of ketosteroid isomerase. *Science* **2014**, *346*, 1510−1514.
(42) Fried, S. D.; Boxer, S. G. Electric field and enzyme catalysis. *Annu. Rev. Biochem.* **2017**, *86*, 387−415.
(43) Anslyn, E. V.; Dougherty, D. A. *Modern Physical Organic Chemistry*, 1st ed.; University Science Books: Sausalito, CA, 2006.
(44) Bard, A. J.; Faulkner, L. R. *Electrochemical Methods: Fundamentals and Applications*, 2nd ed.; Wiley: New York, NY, 2001.
(45) Drobizhev, M.; Molina, R. S.; Callis, P. R.; Scott, J. N.; Lambert, G. G.; Salih, A.; Shaner, N. C.; Hughes, T. E. Local electric field controls fluorescence quantum yield of red and far-red fluorescent proteins. *Front. Mol. Biosci.* **2021**, *8*, 633217.
(46) del Carmen Marín, M.; Agathangelou, D.; Orozco-Gonzalez, Y.; Valentini, A.; Kato, Y.; Abe-Yoshizumi, R.; Kandori, H.; Choi, A.; Jung, K.-H.; Haacke, S.; Olivucci, M. Fluorescence enhancement of a microbial rhodopsin via electronic reprogramming. *J. Am. Chem. Soc.* **2019**, *141*, 262−271.
(47) Romero, P. A.; Arnold, F. H. Exploring protein fitness landscapes by directed evolution. *Nat. Rev. Mol. Cell Biol.* **2009**, *10*, 866−876.
(48) Carrascosa, E.; Bull, J. N.; Scholz, M. S.; Coughlan, N. J. A.; Olsen, S.; Wille, U.; Bieske, E. J. Reversible photoisomerization of the isolated green fluorescent protein chromophore. *J. Phys. Chem. Lett.* **2018**, *9*, 2647−2651.
(49) List, N. H.; Jones, C. M.; Martínez, T. J. Internal conversion of the anionic GFP chromophore: In and out of the I-twisted S_1/S_0 conical intersection seam. *Chem. Sci.* **2022**, *13*, 373−385.
(50) You, M.; Jaffrey, S. R. Structure and mechanism of RNA mimics of green florescent protein. *Annu. Rev. Biophys.* **2015**, *44*, 187−206.
(51) Huang, H.; Suslov, N. B.; Li, N.-S.; Shelke, S. A.; Evans, M. E.; Koldobskaya, Y.; Rice, P. A.; Piccirilli, J. A. A G-quadruplex-containing RNA activates fluorescence in a GFP-like fluorophore. *Nat. Chem. Biol.* **2014**, *10*, 686−691.
(52) Schulten, K.; Hayashi, S. Quantum biology of retinal. In *Quantum Effects in Biology*; Mohseni, M., Omar, Y., Engel, G. S., Plenio, M. B., Eds.; Cambridge University Press: Cambridge, United Kingdom, 2014; pp 237−263.
(53) Guengerich, F. P. Common and uncommon cytochrome P450 reactions related to metabolism and chemical toxicity. *Chem. Res. Toxicol.* **2001**, *14*, 611−650.
(54) Jasniewski, A. J.; Que, L., Jr. Dioxygen activation by nonheme diiron enzymes: Diverse dioxygen adducts, high-valent intermediates, and related model complexes. *Chem. Rev.* **2018**, *118*, 2554−2592.
(55) Nicolet, Y. Structure−function relationships of radical SAM enzymes. *Nat. Catal.* **2020**, *3*, 337−350.
(56) Tantillo, D. J. Importance of inherent substrate reactivity in enzyme-promoted carbocation cyclization/rearrangements. *Angew. Chem., Int. Ed.* **2017**, *56*, 10040−10045.
(57) Khersonsky, O.; Tawfik, D. S. Enzyme promiscuity: A mechanistic and evolutionary perspective. *Annu. Rev. Biochem.* **2010**, *79*, 471−505.
(58) Arnold, F. H. Directed evolution: Bringing new chemistry to life. *Angew. Chem., Int. Ed.* **2018**, *57*, 4143−4148.
(59) Miller, D. C.; Athavale, S. V.; Arnold, F. H. Combining chemistry and protein engineering for new-to-nature biocatalysis. *Nat. Synth.* **2022**, *1*, 18−23.
(60) Neugebauer, M. E.; Kissman, E. N.; Marchand, J. A.; Pelton, J. G.; Sambold, N. A.; Millar, D. C.; Chang, M. C. Y. Reaction pathway engineering converts a radical hydroxylase into a halogenase. *Nat. Chem. Biol.* **2022**, *18*, 171−179.
(61) Mokhtari, D. A.; Appel, M. J.; Fordyce, P. M.; Herschlag, D. High throughput and quantitative enzymology in the genomic era. *Curr. Opin. Struct. Biol.* **2021**, *71*, 259−273.
(62) Tamura, H.; Saito, K.; Ishikita, H. The origin of unidirectional charge separation in photosynthetic reaction centers: nonadiabatic quantum dynamics of exciton and charge in pigment−protein complexes. *Chem. Sci.* **2021**, *12*, 8131−8140.
(63) Weaver, J. B.; Lin, C.-Y.; Faries, K. M.; Mathews, I. I.; Russi, S.; Holten, D.; Kirmaier, C.; Boxer, S. G. Photosynthetic reaction center variants made via genetic code expansion show Tyr at M210 tunes the initial electron transfer mechanism. *Proc. Natl. Acad. Sci. U.S.A.* **2021**, *118*, e211643911.
(64) Mitchell, A. J.; Zhu, Q.; Maggiolo, A. O.; Ananth, N. R.; Hillwig, M. L.; Liu, X.; Boal, A. K. Structural basis for halogenation by iron- and 2-oxo-glutarate-dependent enzyme WelO5. *Nat. Chem. Biol.* **2016**, *12*, 636−640.
(65) Neugebauer, M. E.; Sumida, K. H.; Pelton, J. G.; McMurry, J. L.; Marchand, J. A.; Chang, M. C. Y. A family of radical halogenases for the engineering of amino-acid-based products. *Nat. Chem. Biol.* **2019**, *15*, 1009−1016.
(66) Drienovská, I.; Roelfes, G. Expanding the enzyme universe with genetically encoded unnatural amino acids. *Nat. Catal.* **2020**, *3*, 193−202.
(67) Xiu, X.; Puskar, N. L.; Shanata, J. A. P.; Lester, H. A.; Dougherty, D. A. Nicotine binding to brain receptors requires a strong cation−π interaction. *Nature* **2009**, *458*, 534−537.
(68) Kienhöfer, A.; Kast, P.; Hilvert, D. Selective stabilization of the chorismate mutase transition state by a positively charged hydrogen bond donor. *J. Am. Chem. Soc.* **2003**, *125*, 3206−3207.
(69) Ortmayer, M.; Hardy, F. J.; Quesne, M. G.; Fisher, K.; Levy, C.; Heyes, D. J.; Catlow, C. R. A.; de Visser, S. P.; Rigby, S. E. J.; Hay, S.; Green, A. P. A noncanonical tryptophan analogue reveals an active site hydrogen bond controlling ferryl reactivity in a heme peroxidase. *JACS Au* **2021**, *1*, 913−918.
(70) Lin, C.-Y.; Boxer, S. G. Unusual spectroscopic and electric field sensitivity of chromophores with short hydrogen bonds: GFP and PYP as model systems. *J. Phys. Chem. B* **2020**, *124*, 9513−9525.
(71) Kötzler, M. P.; Robinson, K.; Chen, H.-M.; Okon, M.; McIntosh, L. P.; Withers, S. G. Modulating the nucleophile of a glycoside hydrolase through site-specific incorporation of fluoroglutamic acids. *J. Am. Chem. Soc.* **2018**, *140*, 8268−8276.
(72) Seyedsayamdost, M. R.; Yee, C. S.; Reece, S. Y.; Nocera, D. G.; Stubbe, J. pH rate profiles of F_nY356−R2s (n = 2, 3, 4) in *Escherichia coli* ribonucleotide reductase: Evidence that Y356 is a redox-active amino acid along the radical propagation pathway. *J. Am. Chem. Soc.* **2006**, *128*, 1562−1568.
(73) Wu, Y.; Boxer, S. G. A critical test of the electrostatic contribution to catalysis with non-canonical amino acids in ketosteroid isomerase. *J. Am. Chem. Soc.* **2016**, *138*, 11890−11895.
(74) Torbeev, V. Y.; Hilvert, D. Both the cis-trans equilibrium and isomerization dynamics of a single proline amide modulate β2-microglobulin amyloid assembly. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 20051−20056.