

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

Chi-Yun Lin,<sup>§</sup> Matthew G. Romei,<sup>§</sup> Irimpan I. Mathews, and Steven G. Boxer\*



Cite This: *J. Am. Chem. Soc.* 2022, 144, 3968–3978



Read Online

ACCESS |



Metrics & More

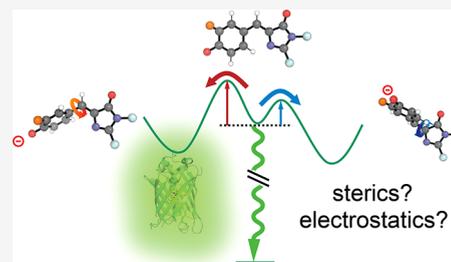


Article Recommendations



Supporting Information

**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.



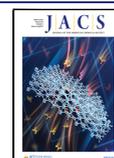
## 1. INTRODUCTION

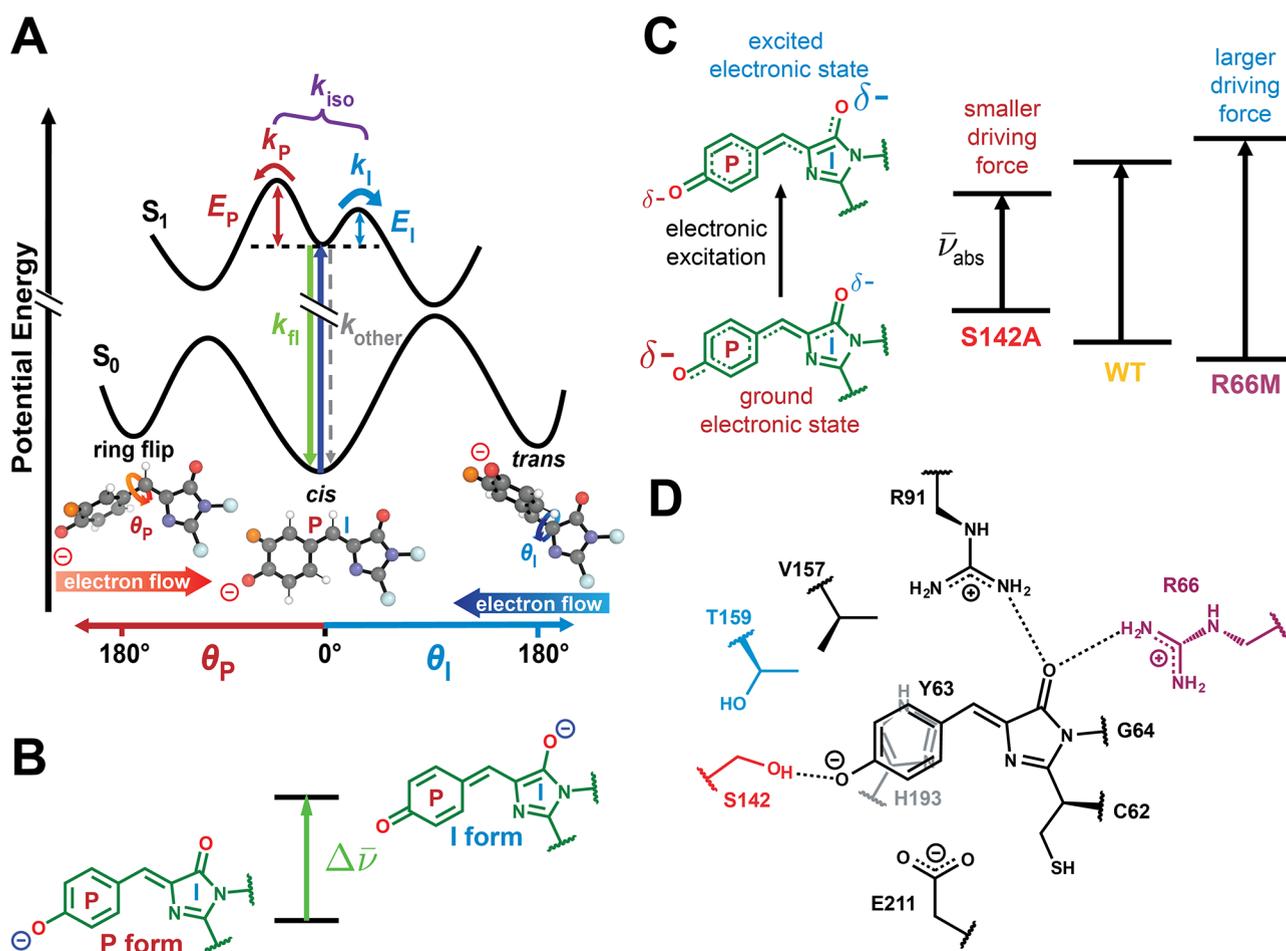
Numerous methods have been employed in developing GFPs with desired behaviors,<sup>1–17</sup> including directed evolution and high-throughput screening of mutant libraries<sup>5–9</sup> that optimize brightness. Machine learning has afforded redder and brighter GFPs,<sup>10,11</sup> and *de novo* protein design has reduced the size of GFP.<sup>12</sup> 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.<sup>13</sup> Photoswitching, the ability to toggle between strongly and weakly fluorescent states through irradiation,<sup>18,19</sup> is another useful function that facilitates super-resolution imaging and optogenetic applications.<sup>20,21</sup> 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<sup>14</sup> and/or painstaking combinations of rational design and screening.<sup>15–17</sup> 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.<sup>22</sup> 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 ( $k_f$  in Figure 1A), the nonmonotonic trend reveals *two* competing nonradiative photoisomerization pathways (Figure 1A) associated with

Received: November 22, 2021

Published: February 24, 2022



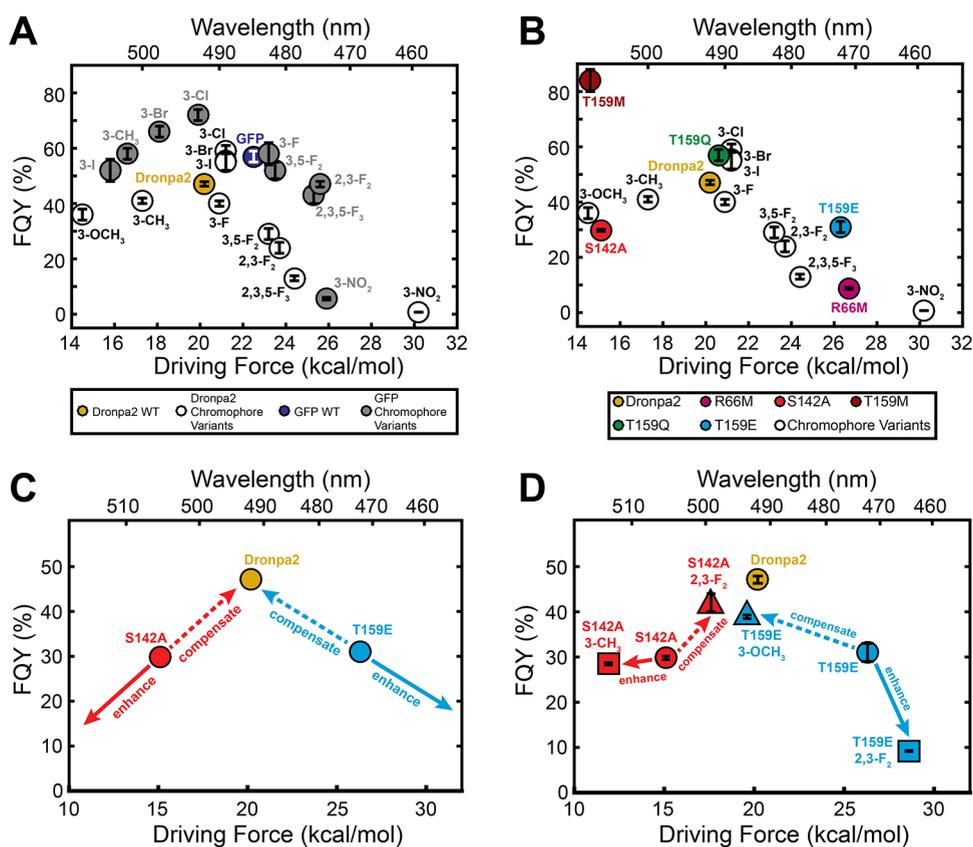


**Figure 1.** Energetics and local environment of the GFP chromophore in Dronpa2. (A) Potential energy surfaces (PESs) for the anionic GFP chromophore along the isomerization coordinate, modified from Figure 4C in ref 22. Copyright (2020) American Association for the Advancement of Science. After excitation from the cis ground state (indigo arrow), the chromophore can either fluoresce ( $k_f$ ) 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,<sup>22</sup> 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,<sup>23</sup> 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  $\Delta\bar{v}$  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,<sup>23</sup> defined as a positive driving force. (C) Marcus–Hush model explaining shifts in transition energy  $\bar{v}_{abs}$  depending on the electrostatic influence of the protein environment on the chromophore’s ground and excited states.<sup>23</sup> (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.<sup>22</sup>

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,<sup>24,25</sup> 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  $\pi$ -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.<sup>24,25</sup> 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.<sup>3,26,27</sup> 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  $\Delta\bar{v}$  (Figure 1B),<sup>23,28</sup> which is the relative energy between the P and I resonance forms of the chromophore, to rigorously quantify electrostatics. We argue that  $\Delta\bar{v}$  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,<sup>23</sup> 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<sup>28</sup> or the more advanced Marcus–Hush treatment.<sup>23</sup> 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  $\bar{\nu}_{\text{abs}}$  (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:<sup>23,28</sup>

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

where  $V_0$  is determined to be  $9530 \text{ cm}^{-1}$ , 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.<sup>23</sup> 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.<sup>23</sup> 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,<sup>23</sup> unify both sources of perturbations,<sup>29</sup> 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<sup>a</sup>**

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

<sup>a</sup>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<sup>-1</sup>. The chromophore modified with OCH<sub>3</sub> possesses a somewhat smaller  $V_0$  than the unsubstituted counterpart,<sup>23</sup> 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<sup>a</sup>**

hybrid protein	predicted combined driving force $\Delta\bar{\nu}$ (cm <sup>-1</sup> )	predicted absorption peak maximum (nm)	observed driving force $\Delta\bar{\nu}$		observed absorption peak maximum (transition energy $\bar{\nu}_{\text{abs}}$ )			FQY (%)	FQY SD (%)
			(cm <sup>-1</sup> )	(kcal/mol)	(cm <sup>-1</sup> )	(nm)	(kcal/mol)		
T159E/2,3-F <sub>2</sub> (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 <sub>3</sub> (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 <sub>2</sub> (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 <sub>3</sub> (enhancing)	4480	510.7	4170	11.9	19510	512.6	55.8	28.5	0.3

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

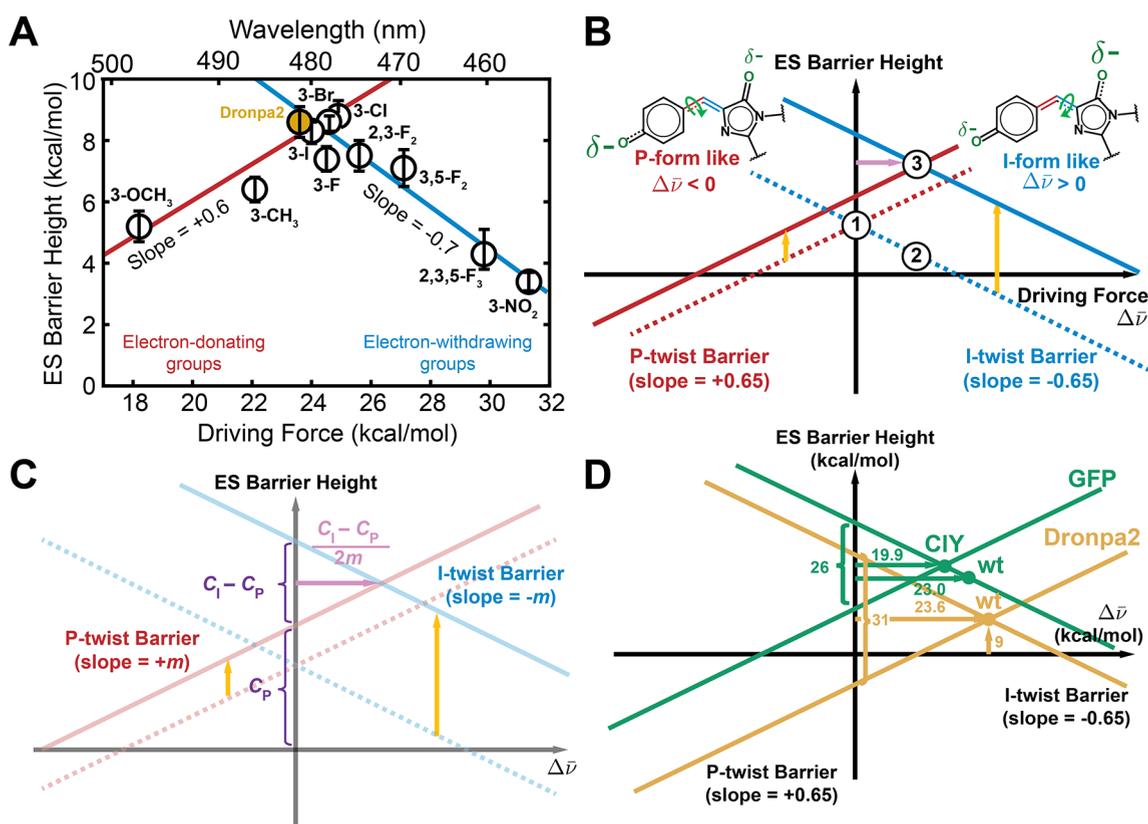
which exhibits a balance between emission and photoisomerization. To isolate the electrostatic effects, residues immediately surrounding the chromophore were replaced with amino acids that minimized differences in size. The S142A mutation causes a red shift by destabilizing the P form through removal 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,<sup>22</sup> 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.<sup>23</sup> For the compensating hybrids, the optimal substituents to bring the driving force of S142A and T159E close to wild type are 2,3-F<sub>2</sub> and 3-OCH<sub>3</sub>, 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<sub>3</sub> and T159E/2,3-F<sub>2</sub>. 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<sub>3</sub> and T159E/2,3-F<sub>2</sub>) 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<sub>2</sub> and T159E/3-OCH<sub>3</sub>) 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<sub>2</sub> or 3-OCH<sub>3</sub>) 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”)<sup>8,30</sup> 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  $\phi_{fl}$  is the ratio between the intrinsic spontaneous emission rate  $k_{fl}$  and the total excited-state decay rate constants<sup>31</sup> (Figure 1A):

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

where  $k_{iso}$  and  $k_{other}$  denote the total rate constant for excited-state isomerization and other nonradiative pathways, respectively;  $\tau$  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;<sup>23,32</sup> 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.  $k_{\text{fl}}$  is estimated to be  $(3.5 \text{ ns})^{-1}$ ,<sup>33</sup> which should be applicable to the unsubstituted GFP chromophore in any environment, so any nonradiative process much slower than this value cannot tune FQYs.

$k_{\text{other}}$  arises from both direct internal conversion and intersystem crossing, but the latter is much less competitive than other excited-state processes<sup>34</sup> (see also Section S2 regarding the substituted chromophores). Accordingly, we can approximate  $k_{\text{other}}$  with a single rate constant from direct internal conversion  $k_{\text{IC}}$  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  $k_{\text{IC}}$ , 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  $k_{\text{iso}}$  uncompetitive; the remaining 23% of excited-state decay can be ascribed to internal conversion;  $k_{\text{IC}}$  is  $(12 \text{ ns})^{-1}$  and imposes an upper limit for GFP's FQY of approximately 80%,<sup>35</sup> 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.<sup>8</sup> 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  $k_{\text{iso}}$  is almost solely responsible for the temperature, electrostatic, and steric dependence of FQY.<sup>22</sup> The associated barriers are typically  $>3$  kcal/mol for GFPs,<sup>22</sup> and the corresponding rate constants are comparable with  $k_{\text{fl}}$  (ns time scale). The rapid intramolecular vibrational energy redistribution (ps time scale)<sup>31,39</sup> 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:

$$k_{\text{iso}} = A_{\text{p}} \exp\left(-\frac{E_{\text{p}}}{RT}\right) + A_{\text{l}} \exp\left(-\frac{E_{\text{l}}}{RT}\right) \approx A \exp\left(-\frac{E}{RT}\right) \quad (3)$$

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

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

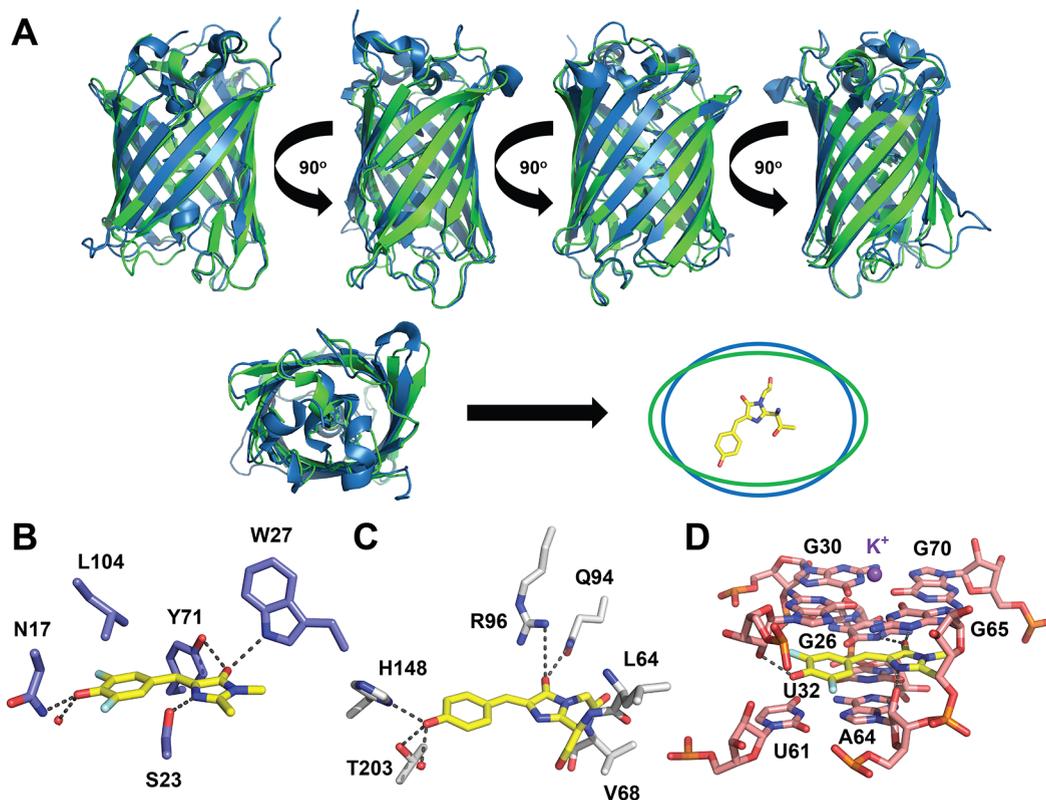
where  $k_{\text{iso}}$  is then approximated with a single Arrhenius expression when we measure the excited-state energy barrier  $E$  of Dronpa2 variants using the temperature dependence of their fluorescence lifetimes.<sup>22</sup>  $A_{\text{p}}$  and  $A_{\text{l}}$  are close in value as experimentally verified,<sup>22,40</sup> so  $A$  should be close to both  $A_{\text{p}}$  and  $A_{\text{l}}$ , and the measured excited-state barrier height  $E$  can well approximate the lesser of the two barriers,  $E_{\text{p}}$  or  $E_{\text{l}}$  (Figure 1A).  $A$  is  $10^3$  to  $10^5 \text{ ns}^{-1}$ ,<sup>22</sup> 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.,  $k_{\text{iso}}$  being 1% of  $k_{\text{fl}}$  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  $E_{\text{p}}$  and  $E_{\text{l}}$ , respectively. The assignment of the P- and I-bond rotation to each value will be explicated later. These slopes are about equal in magnitude ( $\sim 0.65$  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,<sup>41,42</sup> 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:<sup>22</sup>

$$E_{\text{p}} = 0.65\Delta\bar{\nu} + C_{\text{p}} \quad \text{and} \quad E_{\text{l}} = -0.65\Delta\bar{\nu} + C_{\text{l}} \quad (4)$$

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<sup>43</sup> and the Butler–Volmer equation (i.e., linear dependence of free energy barriers on applied potentials) in electrochemistry.<sup>44</sup> 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  $C_{\text{p}}$  and  $C_{\text{l}}$ . We can then rewrite eqs 2 and 3 to explicitly show the electrostatic and steric dependence of the FQY:

force). The electrostatic influence of the red fluorescent protein environment on the corresponding chromophore's FQY is also extensively discussed by a recent paper,<sup>45</sup> while our physical model treats electrostatics differently and

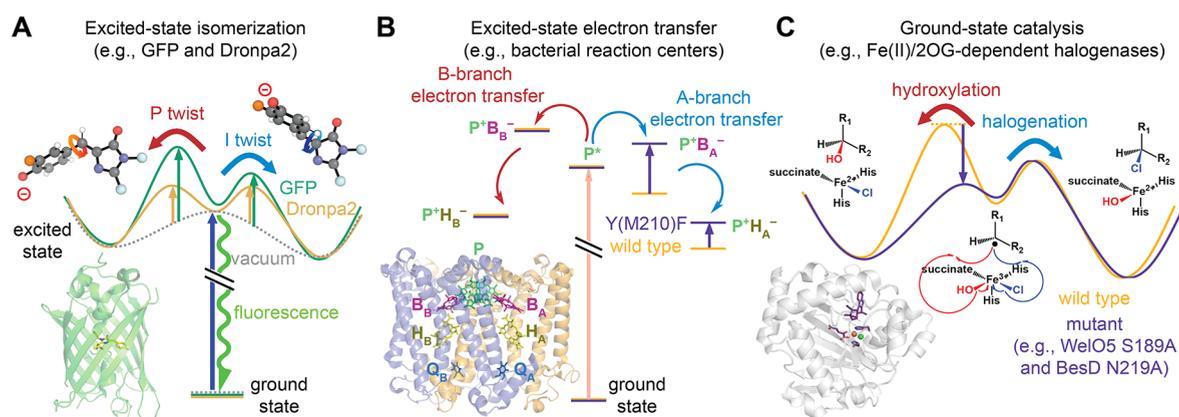


**Figure 4.** GFP chromophore (yellow) in various biomolecular environments. (A) Overlaid  $\beta$ -barrels of Dronpa2 (green, PDB: 6NQJ) and GFP (blue, PDB: 6OFK). The barrels are shown in different perspectives to illustrate the differences in dimensions. The overlaid ovals at the right bottom corner, 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.<sup>46</sup> 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,<sup>30,47</sup> 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).<sup>23</sup> Steric effects  $C_p$  and  $C_l$  serve as an alternative tuning mechanism for the excited-state barriers  $E_p$  and  $E_l$ , preventing the FQY from being completely tied to color via electrostatics, as is the case for other photophysical properties.<sup>23</sup> If  $C_p$  equals  $C_l$ , 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  $C_l$ . 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<sup>23</sup> (Figure 1B). With zero differential sterics from the protein environment ( $C_p = C_l$ ; 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  $C_p = C_l$ , 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.<sup>48</sup> 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.<sup>22,49</sup> If the I ring is anchored inside the protein,  $C_l$  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 *Rhodospirillum rubrum* 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).<sup>62,63</sup> (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).<sup>60,64,65</sup> 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 ( $\sim 50\%$ ; 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 ( $23.6 \times 2 \times 0.65$ , 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%.<sup>22</sup> The corresponding differential sterics, shared by wild-type Dronpa2 and its monochlorinated variant, is 26 kcal/mol ( $= 19.9 \times 2 \times 0.65$ , 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  $\beta$ -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 ( $=$

$(23.6 - 23.0) \times 0.65$ ) 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)<sup>12</sup> 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)<sup>50</sup> to encourage chromophore deprotonation, but fluorines lower the I-twist barrier as electron-withdrawing substituents (Figure 3A).<sup>22</sup> In Spinach,  $\pi$ - $\pi$  stacking with G-quadruplexes effectively inhibits isomerization (Figure 4D),<sup>50,51</sup> 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  $-\text{CF}_3$  group on the I ring and L104H likely explains the much-improved FQY (23%) of chromophore-bound mFAP10.<sup>13</sup>

### 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.<sup>3,27</sup> 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,<sup>52–55</sup> 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,<sup>56</sup> 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 chemistries.<sup>57–60</sup> 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.<sup>61</sup> 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,<sup>66</sup> 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– $\pi$  interactions,<sup>67</sup>  $\pi$ – $\pi$  interactions,<sup>23</sup> hydrogen bonds,<sup>68,69</sup> dipole– $\pi$  interactions,<sup>63</sup>  $pK_a$ ,<sup>70</sup> nucleophilicity,<sup>71</sup> electronic distribution,<sup>22</sup> redox potential,<sup>72</sup> active-site electric field,<sup>73</sup> and isomerization efficiency,<sup>74</sup> can then be quantified by a suitable parameter that is derived from the physical chemistry or physical organic<sup>43</sup> 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.

### ■ ASSOCIATED CONTENT

#### SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/jacs.1c12305>.

Materials and methods and further discussion on intersystem crossing, electrostatic models for fluores-

cence quantum yields for fluorescent proteins, and absolute electrostatic and steric contributions; Additional crystallographic and spectroscopic data in supporting figures and tables (PDF)

### ■ AUTHOR INFORMATION

#### Corresponding Author

Steven G. Boxer – Department of Chemistry, Stanford University, Stanford, California 94305, United States; [orcid.org/0000-0001-9167-4286](https://orcid.org/0000-0001-9167-4286); Email: [sboxer@stanford.edu](mailto: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](https://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](https://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](https://orcid.org/0000-0001-6254-3519)

Complete contact information is available at: <https://pubs.acs.org/10.1021/jacs.1c12305>

#### Author Contributions

<sup>§</sup>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) Pédrelacq, 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  $\beta$ -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.; Verkhusa, 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  $k_{\text{fl}}$  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 fluorescent 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,<sup>36</sup> so  $k_{\text{IC}}$  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,<sup>37,38</sup>  $k_{\text{IC}}$  should be roughly constant across protein environments with  $\beta$ -barrel scaffolds.

- (36) Bochenkova, A. V.; Andersen, L. H. Ultrafast dual photoresponse of isolated biological chromophores: link to the photo-induced 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.; Wolfenden, 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) Ansyn, 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 Marin, 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 fluorescent 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. L.; 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– $\pi$  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) Ortmyer, 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  $F_4Y356-R2s$  ( $n = 2, 3, 4$ ) 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  $\beta$ 2-microglobulin amyloid assembly. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 20051–20056.

Supporting Information of  
**Energetic basis and design of enzyme function demonstrated using GFP, an  
excited-state enzyme**

Chi-Yun Lin<sup>†,§,||</sup>, Matthew G. Romei<sup>†,§,▽</sup>, Irimpan I. Mathews<sup>‡</sup>, Steven G. Boxer<sup>\*,†</sup>

<sup>†</sup>Department of Chemistry, Stanford University, Stanford, CA 94305, USA.

<sup>‡</sup>Stanford Synchrotron Radiation Lightsource, 2575 Sand Hill Road, Menlo Park, CA 94025, USA.

\*To whom correspondence may be addressed. Email: sboxer@stanford.edu.

<sup>§</sup>C.-Y.L. and M.G.R. contributed equally to this work.

<sup>||</sup>Present address: Department of Chemistry, The Pennsylvania State University, University Park, PA 16802, USA.

<sup>▽</sup>Present address: Antibody Engineering, Genentech Inc., South San Francisco, CA 94080, USA.

## SECTION S1: MATERIALS AND METHODS

**Plasmid Construction.** The gene for Dronpa2 (Dronpa-M159T<sup>V2.0</sup>) was generously provided by the Stefan Jakobs lab at the Max Planck Institute for Biophysical Chemistry<sup>1</sup> and inserted into the pET-15b vector (Novagen) at the NdeI and XhoI restriction sites using NEBuilder HiFi DNA assembly (New England Biolabs). Construction of the pET-15b plasmid with the s10:loop:GFP gene (split GFP) was previously described<sup>2</sup>. Point mutations were made using the QuikChange Lightning Site-Directed Mutagenesis Kit (Agilent) according to the manufacturer's protocol. Plasmids containing the appropriate aminoacyl-tRNA synthetase and corresponding tRNA<sup>CUA</sup> for site-specific incorporation of the non-canonical amino acids were described previously<sup>3</sup>.

**DNA Sequence.** Dronpa-M159T<sup>V2.0</sup>  $\equiv$  Dronpa2. Mutation sites are in bold, and the chromophore codons are italicized.

```
ATGGGCAGCAGCCATCATCATCATCACAGCAGCGGCCTGGTGCCGGGCGGCAGCCATATGG
TGAGCAAGGGCGAGGAGAACAACATGGCCGTGATTAAACCAGACATGAAGATCAAGCTGCGTAT
GGAAGGCGCTGTAAATGGACACCCGTTTCGCGATTGAAGGAGTTGGCCTTGGGAAGCCTTTCGAG
GGAAAACAGAGTATGGACCTTAAAGTCAAAGAAGGCGGACCTCTGCCTTTCGCCTATGACATCT
TGACAACTGTGTTCTGTTTACGGCAACAGGGATTCGCCAAATACCCAGAAAATATAGTAGACTA
TTTCAAGCAGTCGTTTCCCTGAGGGCTACTCTTGGGAACGAAGCATGAATTACGAAGACGGGGC
ATTTGTAACGCGACAAACGACATAACCCTGGATGGTGACTGTTATATCTATGAAATTCGATTTG
ATGGTGTGAACTTTCCTGCCAATGGTCCAGTTATGCAGAAGAGGACTGTGAAATGGGAGCCATC
CACTGAGAAATTGTATGTGCGTGATGGAGTGCTGAAGGGTGATGTTAACACGGCTCTGTGCGCTT
GAAGGAGGTGGCCATTACCGATGTGACTTCAAACACTTATAAAGCTAAGAAGGTTGTCCAGT
TGCCAGACTATCACTTTGTGGACCACCACATTGAGATTAAGCCACGACAAAGATTACAGTAA
TGTTAATCTGCATGAGCATGCCGAAGCGCATTCTGGGCTGCCGAGGCAGGCCATGGACGAGCTG
TACAAGTAA
```

**Protein Sequence.** Dronpa-M159T<sup>V2.0</sup>  $\equiv$  Dronpa2. Mutation sites are in bold, and the chromophore forming residues are italicized.

```
MGSSHHHHHSSGLVPGGSHMVSKGEENMAVIKPDMKIKLRMEGAVNGHPFAIEGVGLGKPF
GKQSMDLKVKEGGPLPFAYDILTTFCYGNRVFAKYPENIVDYFKQSFPEGYSWERSMNYEDGG
ICNATNDITLDGDCYIYEIRFDGVNFPANGPVMQKRTVKWEPSTEKLYVRDGVLGKDVNTALS
LEGGHYRCDFKTTYKAKKVVQLPDYHFVDHHEIEIKSHDKDYSNVNLHEHAEAHSGGLPRQAM
DEL
YK
```

**Protein Expression and Purification.** The protein expression and purification protocols for Dronpa2 were performed as previously described<sup>3</sup>. The  $\pi$ - $\pi$  stacking GFP mutants (S65T) were prepared and characterized as in ref. 4. All Dronpa and GFP proteins were stored in a buffer containing 50 mM Tris-HCl at pH 8.0 and 250 mM NaCl at 4°C. The

identity and purity of all proteins were confirmed with electrospray ionization mass spectrometry (ESI-MS) measured with LC-MS (Waters 2795 HPLC with ZQ single quadrupole MS in Stanford University Mass Spectrometry (SUMS) facility). The expected and observed masses are summarized in Table S4.

**UV-Visible Absorption.** PerkinElmer Lambda 25, Lambda 365, and Agilent Cary 6000i UV-Vis spectrometers were used for UV-Vis absorption measurements at room temperature. Data acquisition was performed every 1.0 nm and 0.1 nm at a maximum scan rate of 480 nm/min with the PerkinElmer and Agilent spectrometers, respectively.

**Extinction Coefficient Determination.** Concentrations of the Dronpa point mutants (Y63) were determined by measuring the UV-Vis absorbance at 447 nm in 0.2 M NaOH (Fisher BioReagents, CAS 497-19-8) and scaling by the known extinction coefficient of the deprotonated chromophore in the denatured protein ( $44100 \text{ M}^{-1} \text{ cm}^{-1}$ )<sup>5</sup>. For proteins containing substituted chromophores, the extinction coefficient for the relevant chromophore in the denaturing condition was used as determined previously<sup>3</sup>.

**Fluorescence Quantum Yield Determination.** Absolute fluorescence quantum yields were determined as previously described by comparing sample fluorescence to that of a known standard<sup>6</sup>. Briefly, fluorescence spectra were measured on a PerkinElmer LS 55 fluorescence spectrometer. Excitation and emission slit widths were set to 1.0 nm. The sample was excited at 488 nm, and the emission spectrum was obtained every 0.5 nm at 120 nm/min. Sample absorption was determined using a PerkinElmer Lambda 25 UV-Vis spectrophotometer and then converted into absorption factor according to the following equation:  $f = 1 - 10^{-A_{488}}$ , where  $f$  is the absorption factor and  $A_{488}$  is the absorbance at 488 nm. The absorption factor was plotted against the integrated emission spectrum for different protein concentrations, and the slope of the linear fit was compared against that of fluorescein (Aldrich, CAS 2321-07-5) in 0.1 M NaOH, a known standard with a reported absolute fluorescence quantum yield of 0.90 at 488 nm<sup>7</sup>.

**Protein Crystallization.** Dronpa constructs were crystallized as previously described<sup>3</sup>. Briefly, purified proteins were exchanged into buffer containing 50 mM Tris-HCl at pH 7.8 and 120 mM NaCl using spin centrifugation and concentrated to 10 mg/mL using the extinction coefficients measured in Figure S1. Crystals were grown in the dark at 19°C using the sitting drop method in a 48-well INTELLI-PLATE (Art Robbins Instruments) with

a mother liquor of 0.1 M Tris-HCl at pH 7.4, 0.1 M MgCl<sub>2</sub> (CAS 7786-30-3), and 17.5 – 18.5 % PEG 3350 (Hampton Research, CAS 25322-68-3). The mother liquor was mixed 1:1 with the protein sample to form the initial drop. Looping was performed in limited lighting with minimal time spent in the cryoprotectant (0.1 M Tris-HCl at pH 7.4, 0.1 M MgCl<sub>2</sub>, and 35% PEG 3350) before flash cooling in liquid nitrogen. Any extended time the crystal spent in the cryoprotectant led to crystal cracking and poor diffraction, as discussed further in ref. 3.

**Diffraction Data Collection and Structure Determination.** The x-ray diffraction data for all the Dronpa2 mutants were collected at 100 K at the Stanford Synchrotron Radiation Lightsource (Menlo Park, CA) at BL 12-2. The data were indexed, reduced, and scaled using X-ray Detector Software (XDS)<sup>8,9</sup> using the autoxds script<sup>10</sup>. The structures were solved by molecular replacement with PHENIX<sup>11</sup> using the coordinates from the wild-type room-temperature structure of Dronpa-M159T (PDB: 4UTS)<sup>12</sup>. Chromophore restraint files were built using REEL and eLBOW in PHENIX. Numerous rounds of model building and refinement were performed with Coot<sup>13</sup> and PHENIX. The resulting data collection and refinement statistics are summarized in Table S5.

## SECTION S2: INTERSYSTEM CROSSING FOR SUBSTITUTED CHROMOPHORES

While FPs with the unsubstituted chromophore are found to exhibit negligible intersystem crossing quantum yield ( $\sim 1\%$ )<sup>14</sup>, the substituted chromophores, especially those with -Br, -I, and -NO<sub>2</sub>, known for promoting intersystem crossing in aromatic systems, might show significant singlet-to-triplet conversion. There is no evidence in published data that suggest significant intersystem crossing for these chromophore variants<sup>15,16</sup>. Even though the fluorescence of a YFP-based halide sensor was shown to be quenched by iodide, the external heavy-atom effect was ruled out by testing with similarly bulky anions such as ClO<sub>3</sub><sup>-</sup><sup>17</sup>. In fact, the lack of intersystem crossing even in the presence of heavy atoms are not unheard of in the photochemistry literature. For example, the singlet-to-triplet conversion efficiencies for 3-bromoperylene and perylene are both very low ( $< 1\%$ ), overwhelmed by their efficient fluorescence (FQY  $\sim 100\%$ ) and considerable energy gaps between their singlet and triplet states<sup>18</sup>. The gap between singlet and triplet states for the GFP chromophore is also calculated to be rather large and this likewise discourages intersystem crossing<sup>19</sup>. This is in sharp contrast with molecules known for efficient intersystem crossing, such as anthracene, which has a low-lying T<sub>2</sub> state that is close to S<sub>1</sub><sup>18</sup>. Even with impeded fluorescence emission when substituting Y66 with phenylalanine to yield BFP (FQY  $\sim 1\%$ ), its FQY is not further suppressed by bromination or iodination<sup>20</sup>. Therefore, we conclude that intersystem crossing is negligible across the chromophore variants.

### SECTION S3: SIMILARITIES AND DIFFERENCES IN ELECTROSTATIC MODELS FOR FLUORESCENCE QUANTUM YIELDS OF FLUORESCENT PROTEINS

While our work was in preparation, a paper was published by Drobizhev *et al.* that argued that the electrostatic effect of the red fluorescent protein (RFP) chromophore environment was crucial in modulating the FQY of the corresponding chromophore<sup>21</sup>. There are similarities, but also significant differences between this system and ours, such as our approach to the analysis of these effects and the connection with enzyme design.

The authors prepared a series of RFP environmental mutants (i.e., the chromophore was kept identical) and determined the corresponding spontaneous emission rates and non-radiative decay rates via FQYs and fluorescence lifetimes. While the spontaneous emission rates  $k_{fl}$  were largely consistent (0.1 – 0.2 ns<sup>-1</sup>) across these RFPs and a constant transition dipole moment can be deduced regardless of the protein environment, the nonradiative decay rates showed a wide range between 0.08 – 1.6 ns<sup>-1</sup>, implying that the FQYs are largely determined by the nonradiative decay rates as for GFPs. They then demonstrated that the nonradiative decay rates for 5 out of 7 mutants strongly and monotonically correlated with the corresponding electric fields sensed by the chromophore, as probed through two-photon cross sections. The mutants conforming to the trend were explained with the Marcus formalism, while those deemed outliers were attributed to dominant steric effects or the fluorescing state not being close to the Franck–Condon state. Some design strategies were proposed at the end of the paper based on electrostatic FQY tuning. Interestingly, the photobleaching rate constants for RFP mutants determined by Drobizhev *et al.* nearly a decade ago<sup>22</sup> amounted to 10 – 100 ms<sup>-1</sup> and were also shown to be dependent on protein electrostatics. The differences in nonradiative decay (i.e., mainly excited-state barrier crossing) and photobleaching rates (i.e., reaction after barrier crossing) could originate from a factor of 10<sup>-5</sup> to 10<sup>-4</sup> that partitions chromophore damage vs relaxation that leaves the chromophore intact.

We agree qualitatively with the charge-transfer mechanism of excited-state P twist proposed by Drobizhev *et al.*, and it is satisfying to see the same concept being realized in RFPs. As the I-twist pathway is not adopted due to the presence of the additional acylimine tail attached to the I ring in the RFP chromophore (i.e., the driving force is

completely biased)<sup>23</sup>, there is no excited-state isomerization pathway selection between the P- and I-twists, and thus a monotonic correlation was observed between FQYs and internal electric fields. It is surprising that the mutants in this study are RFPs evolved for applications, such as mFruits, rather than an isosteric series of environmental mutations, yet they still revealed the electrostatic FQY tuning mechanism. The  $\beta$ -barrels of RFPs derived from DsRed<sup>24</sup> must be less tight than those of avGFPs such that steric effects are rarely significant, at least across the studied RFPs. The quantitative interpretations, including the fields experienced by the chromophore and the use of Marcus theory, based on several assumptions that were explicitly presented but are not necessarily correct, need to be addressed. These assumptions can be bypassed by considering models we propose in our current study.

The GFP and RFP chromophores are large probes compared to the length scale of electric field heterogeneity exerted by the protein environment<sup>25</sup>. This very issue has been discussed in our previous paper (see Section S8 in ref. 4). Interpretation of this coarse-grained or “homogeneous” effective field based on local fields, either from calculations or smaller probes, is difficult, and therefore we prefer to use the notion of driving force (i.e., the energy difference of the two underlying valence bond states) to capture the electrostatic influences on the chromophore. The driving force, as stated in the main text, not only is more chemically intuitive, but also encompasses chromophore modifications<sup>4</sup>. The two-state model, in which the driving force is incorporated, can elucidate electrostatic color tuning with fewer parameters than the quadratic Stark effect and exhibits stronger predictive power for other photophysical properties. It can be readily expanded into a three-state model for RFPs by including a new diabatic state with the negative charge localized at the acylimine oxygen<sup>22,23,26</sup>.

It is tempting to invoke the classical (i.e., high-temperature limit) Marcus theory to describe many electron transfer reactions in a wide range of systems, but applying it to excited-state isomerization is questionable. To fit the values from the RFP series and apply Marcus theory, Drobizhev *et al.* had to assume the same reorganization energy for all of these proteins and rely on the harmonic approximation for nuclear degrees of freedom. For outer-sphere electron transfer, this particular approximation holds true because the vibrational energy levels of participating nuclear motions can be treated as

continuous in the high-temperature limit. Even though individual normal modes could be anharmonic, the large number of degrees of freedom renders the central limit theorem applicable, and the Boltzmann distribution from parabolic potentials in one dimension (“the collective coordinate”) is enough to recapitulate the resulting Gaussian population distribution from those normal modes<sup>27,28</sup>. For systems involving one particular high-frequency nuclear mode, such as bond-length alternation (BLA) for the GFP chromophore in the Franck–Condon region<sup>4</sup>, the harmonic approximation is still valid as the motion amplitude is small and anharmonic terms in the diabatic potentials can be safely omitted<sup>29</sup>. A slight adaptation (the Marcus–Hush formalism, generally applied to mixed-valence systems) is required to treat similar systems, which can also flexibly incorporate anharmonic potentials. However, the major reaction coordinates for excited-state isomerization, the bond rotation angles, span a wide range and are periodic, defeating the harmonic approximation, and this is further complicated and made a still worse approximation by sterics from the protein environment. Most simulations of GFP chromophores indicate that BLA and other torsion angles also exhibit nonnegligible changes during isomerization<sup>30–32</sup>. The strong electronic coupling mediated by the exocyclic bonds requires a quantum mechanical treatment and would not predict an inverted region as in weakly coupled donor–acceptor systems but only a monotonic trend. To further complicate the issue, the pre-exponential factor in the Marcus expression, which includes the electronic coupling term, cannot be assumed constant across the series as it is already not a fixed value along the reaction coordinates (especially the bond rotation angles).

A holistic way to address all these problems is to expand the two-state model (or more precisely, valence-bond treatment) into one that includes at least three electronic states (due to the low-lying diradical valence-bond state at large twisting angles<sup>3</sup>) and minimal numbers of strongly coupled nuclear motions, such as twisting angles, one out-of-plane torsion angle, and BLA. The coordinate dependence can be built into the functional form of the electronic coupling. This approach was pioneered by Seth Olsen and Ross McKenzie, who proposed the electronic coupling as a periodic function through Fourier expansion of P- and I-twist angles, and the resulting model even allowed for a conical intersection<sup>33</sup>. The electrostatic effect from the protein environment can then be

incorporated as the relative energy of underlying valence-bond states (i.e., the driving force in the Marcus–Hush model), and the steric influence can be estimated from molecular dynamics simulation of freely rotating chromophores<sup>34</sup> and added directly on top of each electronic level since the sterics should be oblivious to the electronic distribution of the chromophore. Unfortunately, this complicated model requires multiple parameters, such as a coefficient for each term of nuclear coordinate dependence<sup>33</sup>. For the lack of a simpler microscopic model, we opt for a linear approximation for the electrostatic dependence of the excited-state barriers to capture the observed experimental trend and an additive steric component for each isomerization pathway (eq 4), which is sufficient for the observed driving force range for mutants and still affords insights on design principles.

## SECTION S4: ON ABSOLUTE ELECTROSTATIC AND STERIC CONTRIBUTIONS

It would be desirable to experimentally extract the *absolute* electrostatic and steric contributions from the protein environments to excited-state isomerization with respect to the chromophore in vacuum. However, we argue that this is not possible with our current data without further unvalidated assumptions.

First of all, we have to define the corresponding excited-state barrier vs driving force plot for the chromophore in vacuum under electrostatic influences as a reference. The horizontal axis, driving force, is more straightforward. Since the chromophore is not covalently anchored in vacuum as opposed to that in proteins, it should be safe to assume that the exocyclic bonds (P and I bonds) exhibit equal tendencies to isomerize in the excited state if  $\Delta\bar{\nu} = 0$ , so the correlation plot in vacuum should be symmetric with respect to the vertical axis, namely  $\Delta\bar{\nu} = 0$ . However, we should note that the chromophore in vacuum possesses an intrinsic driving force because the chromophore is not symmetric and P ring has a higher electron affinity than I ring. We can estimate this value (24.4 kcal/mol) from the gas-phase absorption maximum of the model GFP chromophore capped with methyl groups (479 nm)<sup>35</sup>, which is close to those of Dronpa2 and GFP. An electric field has to be exerted to negate this intrinsic driving force to truly reach  $\Delta\bar{\nu} = 0$ .

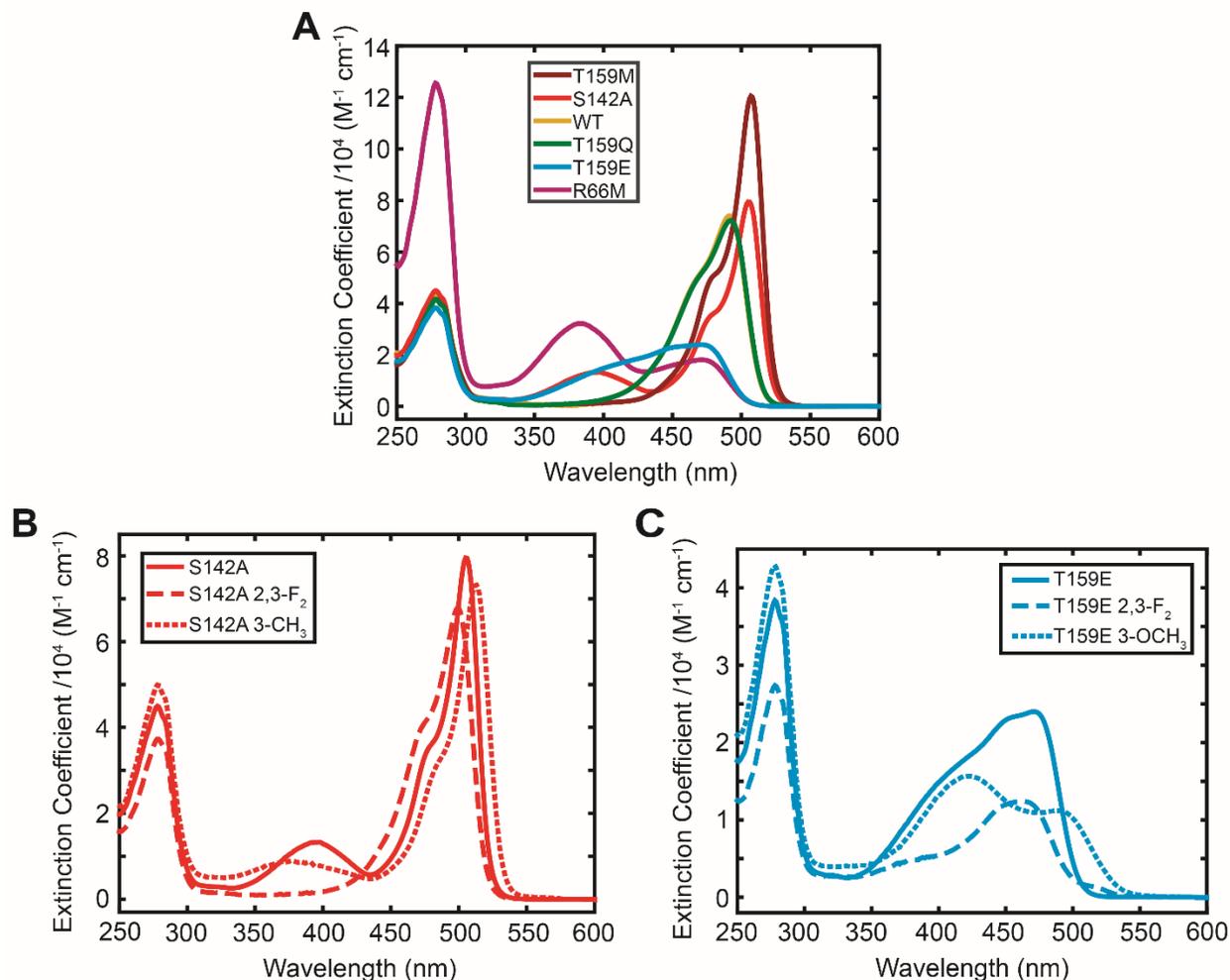
The excited-state barrier, on the other hand, is a much blurrier concept in vacuum. Quantum-mechanical simulations<sup>36,37</sup> and experiments<sup>38</sup> have shown that the anionic chromophore in gas phase is energetically downhill all the way to the conical intersection during I twist in the excited state. To apply eq 4 in the main text, we have implicitly assumed that the nuclear configuration that corresponds to the transition state in the excited state does not change across different environments<sup>39</sup>. In other words, the said nuclear configuration remains to be a local energetic maximum along the isomerization coordinate, which is apparently not true considering the downhill isomerization energetics for the chromophore in vacuum, so we are forced to assign a negative barrier height (which is still reasonable) in this particular situation. But this also means that the “barrier height” would not be experimentally accessible, and therefore we cannot deduce the absolute contributions based on available experimental data. One might wish to assume that the differential sterics (Section 2.3) between the I- and P-bond isomerization can be

solely attributed to the former, such that P-bond isomerization is essentially barrierless<sup>34</sup> and we can thus extract the absolute steric contribution. Yet, this assumption would not be applicable to a wide range of protein environments, especially the GFP  $\pi$ - $\pi$  stacking series and even Dronpa2 itself (with H193).

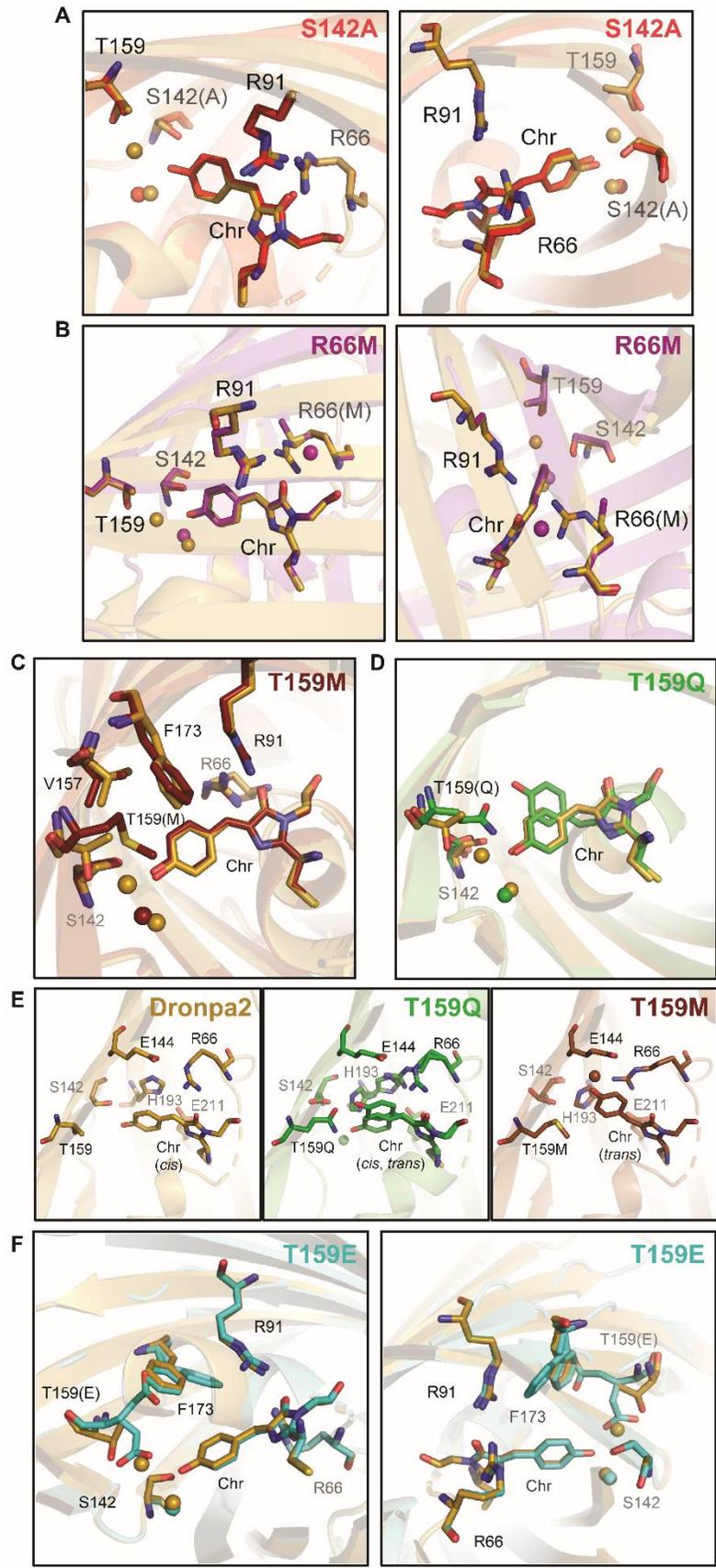
It is also tempting to compare experimental estimates with computational results if such estimation were feasible. We have to caution against such endeavor without addressing the definitions of sterics and electrostatics for each method in the first place. For example, Park and Rhee performed a pioneering theoretical study on the absolute electrostatic and steric contributions of GFP fluorescence, concluding that electrostatics is almost exclusively the cause for the suppression of excited-state isomerization<sup>40</sup>. They defined the steric component as the remaining effect from nullified atomic charges, which is theoretically practical but experimentally impossible. We, in contrast, use the transition energy as a metric for electrostatic contribution. In other words, we compare the excited-state barriers against the electrostatic environment experienced by the chromophore before and right after excitation, assuming that there is no electronic polarization involved in the protein environment, to extract the electrostatic contribution to isomerization. This experimental approach may seem straightforward, but it implies that any *change in the electrostatic environment* during excited-state isomerization with respect to what is experienced by the chromophore with the Franck–Condon geometry is counted towards *sterics*. For example, for the  $\pi$ - $\pi$  stacking GFP mutants, the P ring experiences a considerable change in the  $\pi$ - $\pi$  stacking interaction no matter which isomerization pathway is taken. In Park and Rhee’s study, this aspect would have been attributed to electrostatics, while for us this is a combination of electrostatics (due to its ability to tune the color) and sterics (changed interaction during isomerization). Similarly, a loss of hydrogen bond to S142 in Dronpa2 during I twist is partially viewed as a steric obstruction to isomerization with our method, which can quickly turn into a fiery semantic debate. The simulation strategy adopted by Chen *et al.*<sup>34</sup> would be closer to our classification, since the charges are not nullified but instead fixed at ground-state values, and both exocyclic bonds are allowed to freely rotate to sample the steric barrier exerted by the protein environment. Since sterics is naturally short-ranged electrostatics (among all four fundamental interactions)<sup>41,42</sup>, the dissection of sterics and electrostatics are merely

artifacts to facilitate human understanding, so is the classification of different intermolecular forces from chemists. Because we still have to translate electrostatics and sterics into atomic positions when bringing protein design to fruition, so long as a framework can bear intuition, offer predictions, and infer design principles, such framework should be considered useful.

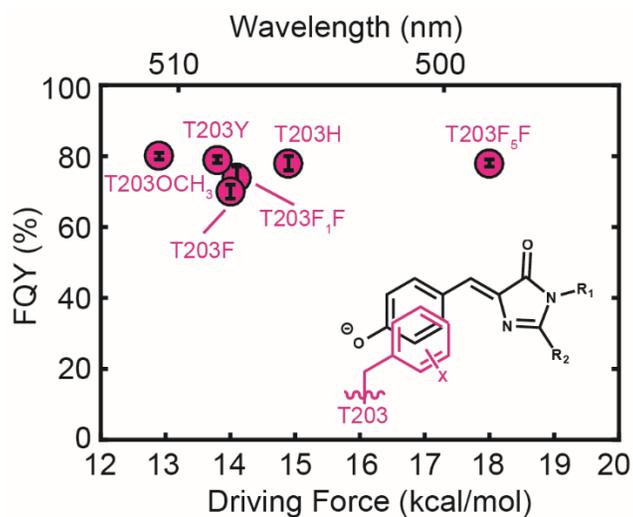
## SECTION S5: SUPPORTING FIGURES



**Figure S1. Absorption spectra of the Dronpa2 mutants and hybrids normalized to their extinction coefficients.** (A) Mutants. The large peak at 280 nm for the R66M mutant is due to the high proportion of protein with immature chromophores due to the hindered chromophore maturation caused by the mutation. Immature chromophores do not absorb in the visible region. (B) Hybrids with S142A mutation. (C) Hybrids with T159E mutation. Transition energies for the anionic chromophore are determined from the reddest absorption maxima.



**Figure S2. Comparison between the chromophore environments of WT Dronpa2 (gold; PDB: 6NQJ) and S142A, R66M mutants, or T159 isosteric mutant series.** (A) In the S142A mutant (red; PDB: 7RRI), the hydrogen bond partner to the phenolate oxygen, S142, is lost, and no clear replacement can be found. Loss of the interaction with S142 slightly impacts the chromophore's P-bond and I-bond dihedral angles by between 10° and 15°, but the corresponding photophysical properties should be minimally affected based on the angular overlap between the p orbitals forming the  $\pi$  orbital at the P and I bonds (see Equation S53 of ref. 4). (B) In the R66M mutant (magenta; PDB: 7RRH), a water molecule replaces R66 as a hydrogen bond partner of the chromophore's imidazolinone oxygen as the side chain of methionine is shorter than that of arginine. Despite the retention of a hydrogen bond partner, the nature of the interaction has changed as a positive charge is removed. Different perspectives are shown in the left and right panels. (C) In the T159M mutant (maroon; PDB: 2IOV), the methionine displaces a water molecule, which is hydrogen bonded to the phenolate oxygen in the WT structure. Other than residues V157 and F173, which are nonpolar and rearrange slightly to accommodate the structural perturbation, the rest of the residues within the chromophore pocket, as well as chromophore dihedral angles, remain largely unchanged from those of Dronpa2. (D) In the T159Q mutant (green; PDB: 7RRJ), the chromophore exists in both the *cis* (~75% occupancy) and *trans* (~25% occupancy) conformations. The glutamine side chain hydrogen bonds to the phenolate oxygen in both conformations, replacing S142 when in the *cis* conformation. (E) Comparison between the chromophore environments of the T159Q mutant (green, center) and the typical *cis* (gold, left) and *trans* (brown, right) conformations. Note that the *trans* conformation on the right is from the T159M mutant (PDB: 2POX), since no photoisomerized structures have been solved of Dronpa2. The chromophore's conformation is linked to the orientations for the side chains of R66, H193, and S142. (F) In the T159E mutant (blue; PDB: 7RRK), E159 and F173 exist in two conformations within the chromophore pocket. In one conformation (60% occupancy), the glutamate replaces a water molecule as a hydrogen bond partner to the phenolate oxygen and F173 adopts the same conformation as in WT. In the other conformation (40% occupancy), the glutamate adopts a different rotamer pointing away from the chromophore, forcing F173 to rotate towards R91. These two conformations likely represent the two protonation states of the chromophore: the deprotonated (anionic) B state and the protonated (neutral) A state (Figure S1A).



**Figure S3.** FQY plotted against driving force for GFP with substituted and unsubstituted aromatic ring side chains at residue 203 (Table S2). The transition energy, from which the driving force is calculated (eq 1), is obtained from ref. 4. This demonstrates the insensitivity of FQY despite the large change in driving force. F<sub>5</sub>F: 2,3,4,5,6-pentafluorophenylalanine; F<sub>1</sub>F: 4-fluorophenylalanine.

## SECTION S6: SUPPORTING TABLES

**Table S1. Driving forces, absorption peak maxima, and fluorescence quantum yields (FQYs) for each Dronpa2 point mutant.** Driving force is calculated from eq 1 in the main text with an electronic coupling  $V_0$  of  $9530 \text{ cm}^{-1}$ . SD: standard deviation.

Environmental Mutant	Driving Force $\Delta\bar{\nu}$		Absorption Peak Maximum (Transition Energy $\bar{\nu}_{abs}$ )			FQY (%)	FQY SD (%)
	( $\text{cm}^{-1}$ )	(kcal/mol)	( $\text{cm}^{-1}$ )	(nm)	(kcal/mol)		
R66M	9350	26.7	21230	471.0	60.7	8.8	0.4
T159E	9200	26.3	21160	472.5	60.5	31	2
T159Q	7200	20.6	20370	490.8	58.3	57	2
Dronpa2 ("wild type")	7010	20.2	20310	492.4	58.1	46	2
S142A	5300	15.1	19780	505.5	56.6	29.8	0.4
T159M	5100	14.6	19730	506.8	56.4	84	4

**Table S2. Driving forces, absorption peak maxima, and fluorescence quantum yields (FQYs) for GFP with substituted and unsubstituted aromatic ring side chains at residue 203.** SD: standard deviation; F<sub>5</sub>F: 2,3,4,5,6-pentafluorophenylalanine; F<sub>1</sub>F: 4-fluorophenylalanine.

Environmental Mutant	Driving Force $\Delta\bar{\nu}$		Absorption Peak Maximum (Transition Energy $\bar{\nu}_{abs}$ )			FQY (%)	FQY SD (%)
	(cm <sup>-1</sup> )	(kcal/mol)	(cm <sup>-1</sup> )	(nm)	(kcal/mol)		
T203F <sub>5</sub> F	6310	18.0	20080	498.1	57.4	78	1
T203H	5220	14.9	19760	506.0	56.5	78	2
T203F <sub>1</sub> F	4940	14.1	19690	507.9	56.3	74	3
T203F	4910	14.0	19680	508.1	56.3	70	2
T203Y	4840	13.8	19670	508.5	56.2	79	1
T203OCH <sub>3</sub> Y	4500	12.9	19580	510.6	56.0	80	1

**Table S3. Driving forces, absorption maxima, fluorescence quantum yields (FQYs), and excited-state barriers for Dronpa2 and GFP chromophore variants.** The variants are named by the chromophore substituents. The values are extracted from ref. 3. While the thermochromic shifts are small for GFP variants, those for Dronpa2 variants are significant (~ 10 nm) but consistent across the series (Figure S6 in ref. 3). Figures 2 and 3 are plotted with the 293-K and 77-K driving forces, respectively. Bolded values are used in Section 2.3 and Figure 3D for estimating steric and electrostatic contributions.

Dronpa2 variants	Driving Force $\Delta\bar{\nu}$ (kcal/mol)		Absorption Peak Maximum (nm)		FQY (%)	Excited-state barrier (kcal/mol)
	293 K	77 K	293 K	77 K		
3-NO <sub>2</sub>	30.2	31.3	459	454.8	0.73	3.4
2,3,5-F <sub>3</sub>	24.4	29.8	479	460.5	13.0	4.3
3,5-F <sub>2</sub>	23.1	27.1	483	469.8	29	7.1
2,3-F <sub>2</sub>	23.7	25.6	481	474.9	24	7.5
3-F	20.9	24.5	490	478.5	40	7.3
3-Cl	21.2	24.9	489	477.3	59	8.8
3-Br	21.2	24.6	489	478.2	55.6	8.6
3-I	21.2	24.0	489	480.3	55	8.3
WT	20.2	<b>23.6</b>	492	481.5	<b>47.1</b>	<b>8.6</b>
3-CH <sub>3</sub>	17.3	22.1	500	486.3	41	6.4
3-OCH <sub>3</sub>	14.5	18.2	507	497.7	36	5.2
GFP variants	Driving Force $\Delta\bar{\nu}$ (kcal/mol)		Absorption Peak Maximum (nm)		FQY (%)	Excited-state barrier (kcal/mol)
	293 K	77 K	293 K	77 K		
3-NO <sub>2</sub>	25.9	26.7	474	471.0	5.6	Not Determined
2,3,5-F <sub>3</sub>	25.3	24.6	476	478.2	43	
3,5-F <sub>2</sub>	23.4	23.9	482	480.6	52	
2,3-F <sub>2</sub>	25.6	24.6	475	478.2	47	
3-F	23.1	23.6	483	481.5	58	
3-Cl	19.8	<b>19.9</b>	493	492.9	<b>72</b>	
3-Br	18.1	18.3	498	497.4	66	
3-I	15.8	15.3	504	505.2	52	
WT	22.5	<b>23.0</b>	485	483.3	57	
3-CH <sub>3</sub>	16.6	16.5	502	502.2	58	
3-OCH <sub>3</sub>	9.4	7.4	517	519.9	11.4	

**Table S4. Expected and measured mass for each Dronpa construct.**

Protein Constructs	Expected Mass <sup>a</sup> (Da)	Observed Mass <sup>b</sup> (Da)
Dronpa	28947	28953
Dronpa2 (M159T)	28917	28920
Dronpa2 T159Q	28944	28950
Dronpa2 T159E	28945	28950
Dronpa2 S142A	28901	28907
Dronpa2 R66M	28892	28917 <sup>c</sup>
Dronpa2 T159E Y63(2,3-F <sub>2</sub> Y)	28981	28986
Dronpa2 T159E Y63(3-OCH <sub>3</sub> Y)	28977	28983
Dronpa2 S142A Y63(2,3-F <sub>2</sub> Y)	28937	28944
Dronpa2 S142A Y63(3-CH <sub>3</sub> Y)	28915	28919

<sup>a</sup> Predicted from the primary sequence with N-terminal methionine removed.

<sup>b</sup> Proteins with ~ 30 kDa have  $\pm 10$  Da deviations, depending on the protonation states.

<sup>c</sup> The observed mass corresponds to the protein with an immature chromophore, which adds 20 Da to the expected mass due to oxidation and dehydration reactions.

**Table S5. X-ray data collection and refinement statistics for Dronpa2 mutants.**

	<b>Dronpa2 R66M</b>	<b>Dronpa2 S142A</b>	<b>Dronpa2 T159Q</b>	<b>Dronpa2 T159E</b>
PDB entry	<b>7RRH</b>	<b>7RRI</b>	<b>7RRJ</b>	<b>7RRK</b>
<b>Data collection statistics</b>				
beamline	BL12-2			
wavelength (Å)	0.9795			
detector distance (mm)	300	500	400	370
resolution range (Å)	36.05 – 1.747 (1.81 – 1.747)	37.29 – 2.643 (2.738 – 2.643)	36.89 – 2.198 (2.277 – 2.198)	37.73 – 1.929 (1.998 – 1.929)
space group	C 1 2 1 (No.5)	P 1 2 <sub>1</sub> 1 (No.4)		
unit cell dimensions a, b, c (Å) $\alpha, \beta, \gamma$ (°)	186.63, 71.54, 108.27 90.0, 110.12, 90.0	79.41, 85.81, 143.17 90.0, 95.0, 90.0	79.75, 85.79, 144.05 90.0, 95.14, 90.0	79.83, 85.73, 143.05 90.0, 95.0, 90.0
total observations	1788832 (161616)	723035 (66493)	1203995 (109488)	1738413 (154307)
unique reflections	132972 (12539)	54577 (5081)	96708 (6925)	139109 (13363)
multiplicity	13.5 (12.7)	13.2 (12.8)	12.4 (11.5)	12.5 (11.5)
completeness (%)	97.9 (92.9)	94.7 (90.8)	93.3 (68.5)	95.9 (92.7)
mean I/ $\sigma$ I	18.3 (1.2)	7.0 (1.2)	10.5 (1.2)	10.8 (1.2)
Wilson B-factor (Å <sup>2</sup> )	32.3	34.3	30.3	27.1
R <sub>merge</sub>	0.069 (1.87)	0.277 (1.60)	0.268 (1.86)	0.135 (1.70)
R <sub>meas</sub>	0.071 (1.94)	0.289 (1.67)	0.279 (1.95)	0.140 (1.77)
R <sub>pim</sub>	0.019 (0.54)	0.079 (0.46)	0.078 (0.57)	0.039 (0.50)
CC <sub>1/2</sub>	1 (0.703)	0.995 (0.747)	0.997 (0.837)	0.998 (0.735)
CC*	1	0.999	0.999	1

	(0.908)	(0.925)	(0.955)	(0.920)
<b>Refinement statistics</b>				
reflections	132788	53355	91981	138654
used	(12531)	(5074)	(6680)	(13345)
reflections	6641	2671	4600	6927
used for R <sub>free</sub>	(627)	(255)	(337)	(667)
R <sub>work</sub>	0.1773	0.2256	0.2409	0.2439
R <sub>free</sub>	0.2006	0.2767	0.2681	0.2666
number of				
non-H atoms	11331	14128	15127	15363
protein	10344	13703	14010	13942
ligand	126	168	273	168
solvent	861	257	844	1253
protein residues	1290	1743	1752	1727
RMSD bond				
lengths (Å)	0.010	0.013	0.011	0.011
RMSD bond				
angles (°)	1.34	1.44	1.29	1.35
Ramachandran				
favored (%)	99.05	97.88	98.83	98.75
Ramachandran				
allowed (%)	0.95	2.06	1.00	1.25
Ramachandran				
outliers (%)	0.00	0.06	0.18	0.00
rotamer outliers				
(%)	1.25	3.88	2.20	2.13
clashscore	3.42	9.61	10.49	9.10
Average				
B-factor (Å <sup>2</sup> )	47.92	38.52	38.07	43.57
protein	48.07	38.75	38.16	44.00
ligand	46.40	34.41	36.41	35.73
solvent	46.38	29.05	37.10	39.74

## SECTION S7: REFERENCES

- [1] Willig, K. I.; Stiel, A. C.; Brakemann, T.; Jakobs, S.; Hell, S. W. Dual-label STED nanoscopy of living cells using photochromism. *Nano Lett.* **2011**, *11*, 3970–3973.
- [2] Lin, C.-Y.; Both, J.; Do, K.; Boxer, S. G. Mechanism and bottlenecks in strand photodissociation of split green fluorescent proteins (GFPs). *Proc. Natl. Acad. Sci. USA* **2017**, *114*, E2146–E2155.
- [3] Romei, M. G.; Lin, C.-Y.; Mathews, I. I.; Boxer, S. G. Electrostatic control of photoisomerization pathways in proteins. *Science* **2020**, *367*, 76–79.
- [4] 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.
- [5] Ward, W. W. Properties of the Coelenterate Green-Fluorescent Proteins. In *Bioluminescence and Chemiluminescence*; DeLuca, M. A., McElroy, W. D., Eds.; Academic Press, 1981; pp 235–242.
- [6] Würth, C.; Grabolle, M.; Pauli, J.; Spieles, M.; Resch-Genger, U. Relative and absolute determination of fluorescence quantum yields of transparent samples. *Nat. Protoc.* **2013**, *8*, 1535–1550.
- [7] Gaigalas, A. K.; Wang, L. Measurement of the fluorescence quantum yield using a spectrometer with an integrating sphere detector. *J. Res. Natl. Inst. Stand. Technol.* **2008**, *113*, 17–28.
- [8] Kabsch, W. XDS. *Acta Cryst.* **2010**, *D66*, 125–132.
- [9] Kabsch, W. Integration, scaling, space-group assignment and post-refinement. *Acta Cryst.* **2010**, *D66*, 133–144.
- [10] Using the *autoxds* script.  
[http://smb.slac.stanford.edu/facilities/software/xds/#autoxds\\_script](http://smb.slac.stanford.edu/facilities/software/xds/#autoxds_script) (accessed Nov 16, 2021).
- [11] Adams, P. D.; Afonine, P. V.; Bunkóczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L.-W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H. PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Cryst.* **2010**, *D66*, 213–221.

- [12] Kaucikas, M.; Fitzpatrick, A.; Bryan, E.; Struve, A.; Henning, R.; Kosheleva, I.; Srajer, V.; Groenhof, G.; van Thor, J. J. Room temperature crystal structure of the fast switching M159T mutant of the fluorescent protein Dronpa. *Proteins* **2015**, *83*, 397–402.
- [13] Emsley, P.; Cowtan, K. Coot: Model-building tools for molecular graphics. *Acta Cryst.* **2004**, *D60*, 2126–2132.
- [14] 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.
- [15] Oscar, B. G.; Zhu, L.; Wolfenden, 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.
- [16] Boulanger, S. A.; Chen, C.; Tang, L.; Zhu, L.; Baleeva, N. S.; Myasnyanko, I. N.; Baranov, M. S.; Fang, C. Shedding light on ultrafast ring-twisting pathways of halogenated GFP chromophores from the excited to ground state. *Phys. Chem. Chem. Phys.* **2021**, *23*, 14636–14648.
- [17] Jayaraman, S.; Haggie, P.; Wachter, R. M.; Remington, S. J.; Verkman, A. S. Mechanism and cellular applications of a green fluorescent protein-based halide sensor. *J. Mol. Biol.* **2000**, *275*, 6047–6050.
- [18] Turro, N. J.; Ramamurthy, V.; Scaiano, J. C. *Modern Molecular Photochemistry of Organic Molecules*, 1st ed.; University Science Books: Sausalito, California, 2010.
- [19] Bravaya, K. B.; Khrenova, M. G.; Grigorenko, B. L.; Nemukhin, A. V.; Krylov, A. I. Effect of protein environment on electronically excited and ionized states of the green fluorescent protein chromophore. *J. Phys. Chem. B* **2011**, *115*, 8296–8303.
- [20] Wang, L.; Xie, J.; Deniz, A. A.; Schultz, P. G. Unnatural amino acid mutagenesis of green fluorescent protein. *J. Org. Chem.* **2003**, *68*, 174–176.
- [21] 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.

- [22] 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.
- [23] Olsen, S.; Smith, S. C. Radiationless decay of red fluorescent protein chromophore models via twisted intramolecular charge-transfer states. *J. Am. Chem. Soc.* **2007**, *129*, 2054–2065.
- [24] Shaner, N. C. The mFruit collection of monomeric fluorescent proteins. *Clin. Chem.* **2013**, *59*, 440–441.
- [25] Park, J. W.; Rhee, Y. M. Electric field keeps chromophore planar and produces high yield fluorescence in green fluorescent protein. *J. Am. Chem. Soc.* **2016**, *138*, 13619–13629.
- [26] Chica, R. A.; Moore, M. M.; Allen, B. D.; Mayo, S. L. Generation of longer emission wavelength red fluorescent proteins using computationally designed libraries. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 20257–20262.
- [27] Chandler, D. Electron Transfer in Water and Other Polar Environments, How It Happens. In *Classical and Quantum Dynamics in Condensed Phase Simulations*; Berne, B. J., Ciccotti, G., Coker, D. F., Eds.; World Scientific, 1998; pp 25–49.
- [28] Peters, B. Common features of extraordinary rate theories. *J. Phys. Chem. B* **2015**, *119*, 6349–6356.
- [29] Blumberger, J. Recent advances in the theory and molecular simulation of biological electron transfer reactions. *Chem. Rev.* **2015**, *115*, 11191–11238.
- [30] Gozem, S.; Luk, H. L.; Schapiro, I.; Olivucci, M. Theory and simulation of the ultrafast double-bond isomerization of biological chromophores. *Chem. Rev.* **2017**, *22*, 13502–13565.
- [31] Virshup, A. M.; Punwong, C.; Pogorelov, T. V.; Lindquist, B. A.; Ko, C.; Martínez, T. J. Photodynamics in complex environments: Ab initio multiple spawning quantum mechanical/molecular mechanical dynamics. *J. Phys. Chem. B* **2009**, *113*, 3280–3291.

- [32] Gromov, E. V.; Domratcheva, T. Four resonance structures elucidate double-bond isomerisation of a biological chromophore. *Phys. Chem. Chem. Phys.* **2020**, *22*, 8535–8544.
- [33] Olsen, S.; McKenzie, R. H. A diabatic three-state representation of photoisomerization in the green fluorescent protein chromophore. *J. Chem. Phys.* **2009**, *130*, 184302.
- [34] Chen, M. C.; Lambert, C. R.; Urgitis, J. D.; Zimmer, M. Photoisomerization of green fluorescent protein and the dimensions of the chromophore cavity. *Chem. Phys.* **2001**, *270*, 157–164.
- [35] Nielsen, S. B.; Lapierre, A.; Andersen, J. U.; Pedersen, U. V.; Tomita, S.; Andersen, L. H. Absorption spectrum of the green fluorescent protein chromophore anion *in vacuo*. *Phys. Rev. Lett.* **2001**, *87*, 228102.
- [36] 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.
- [37] 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.
- [38] Carrascoasa, 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.
- [39] Fried, S. D.; Boxer, S. G. Electric field and enzyme catalysis. *Annu. Rev. Biochem.* **2017**, *86*, 387–415.
- [40] Park, J. W.; Rhee, Y. M. Electric field keeps chromophore planar and produces high yield fluorescence in green fluorescent protein. *J. Am. Chem. Soc.* **2016**, *138*, 13619–13629.
- [41] Anslyn, E. V.; Dougherty, D. A. *Modern Physical Organic Chemistry*, 1st ed.; University Science Books: Sausalito, California, 2006.
- [42] Laughlin, R. B.; Pines, D. The theory of everything. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 28–31.