ARTICLE

Structural origin and rational development of bright red noncanonical variants of green fluorescent protein

Cheng Chen,[a] Hao Zhang,[bc] Jing Zhang,[c] Hui-wang Ai,*[bcd] and Chong Fang*[a]

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

The incorporation of noncanonical amino acids (ncAAs) into fluorescent proteins is promising for red-shifting their fluorescence and benefiting tissue imaging with deep penetration and low phototoxicity. However, ncAA-based red fluorescent proteins (RFPs) have been rare. The 3-aminotyrosine modified superfolder green fluorescent protein (aY-sfGFP) represents a recent advance, yet the molecular mechanism for its red-shifted fluorescence remains elusive while its dim fluorescence hinders applications. Herein, we implement femtosecond stimulated Raman spectroscopy to obtain structural fingerprints in the electronic ground state and reveal that aY-sfGFP possesses a GFP-like instead of RFP-like chromophore. Red color of aY-sfGFP intrinsically arises from a unique "double-donor" chromophore structure that raises ground-state energy and enhances charge transfer, notably differing from the conventional conjugation mechanism. We further developed two aY-sfGFP mutants (E222H and T203H) with significantly improved (~12-fold higher) brightness by rationally restraining the chromophore's nonradiative decay through electronic and steric effects, aided by solvatochromic and fluorogenic studies of the model chromophore in solution. This study thus provides functional mechanisms and generalizable insights into ncAA-RFPs with an efficient route for engineering redder and brighter fluorescent proteins.

1. Introduction

One key motto for bioimaging, redder is better, demands a versatile set of noninvasive photosensitive biomarkers. Red fluorescent proteins (RFPs) are powerful tools for fluorescence imaging due to reduced tissue absorption, autofluorescence, and phototoxicity at long wavelengths. Therefore, the targeted engineering of bright RFPs is highly desirable. Following the development of *Aequorea victoria* green fluorescent protein (avGFP) variants spanning the blue to yellow spectral region, a large family of RFPs with fluorescence emission over 550 nm has been discovered and developed from the *Anthozoa* species.[1-5] However, the existing RFPs commonly display shortcomings in their brightness, maturation, oligomerization, stability, and dynamic range in biosensing.[2-4] Many efforts were devoted to converting GFP to RFPs but have resulted in limited success.[6-8]

Several studies addressed these issues by incorporating noncanonical amino acids (ncAAs) into GFP.[9,10] Among these efforts, tyrosine-66 as part of the GFP chromophore was typically replaced with ncAAs to generate the modified chromophores, yielding emission redshifts.[11,12] It was recently discovered that introducing 3-aminotyrosine (aY) into the GFP chromophore (Fig. 1) can give rise to an impressive ~100 nm redshift in fluorescence and effectively turn it into an RFP.[13,14]

This approach can be generalized to diverse GFP variants and biosensors, and further applied to *in vitro* assays and live-

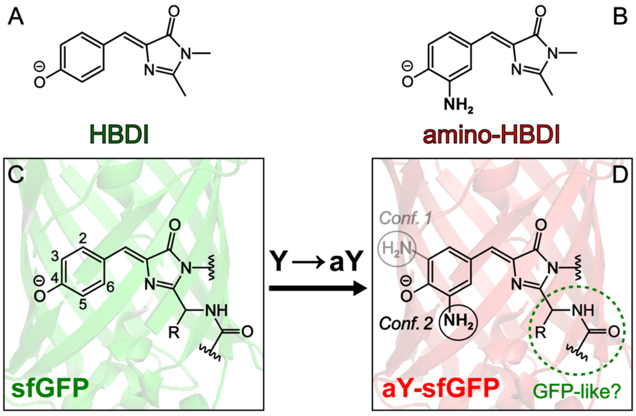

Fig. 1 The green-to-red conversion in superfolder GFP (sfGFP) by an amino substitution. The model chromophores (A) HBDI and (B) amino-HBDI in the deprotonated state are directly compared with (C) sfGFP with a deprotonated chromophore and (D) aY-sfGFP with an amino-substituted deprotonated chromophore, respectively. Two possible conformations with the singly substituted amino group in aY-sfGFP are indicated by gray and black solid circles. The tentative structural moiety to be examined in this work is highlighted by the green dashed ellipse.

---

[a] Department of Chemistry, Oregon State University, 153 Gilbert Hall, Corvallis, Oregon 97331, USA.
E-mail: Chong.Fang@oregonstate.edu; Web: https://fanglab.oregonstate.edu/
[b] Department of Chemistry, University of Virginia, Charlottesville, Virginia 22904, USA.
E-mail: Huiwang.Ai@virginia.edu
[c] Department of Molecular Physiology and Biological Physics and Center for Membrane and Cell Physiology, University of Virginia School of Medicine, Charlottesville, Virginia 22908, USA.
[d] The UVA Comprehensive Cancer Center, University of Virginia, Charlottesville, Virginia 22908, USA.

† Electronic Supplementary Information (ESI) available: Supplementary analysis and discussions (Sections S1.1–1.5) on the solvatochromism of amino-HBDI and various photophysical characterization of model chromophores and FPs, supplementary Fig. S1–S14 and Tables S1–S6 on steady-state and time-resolved spectral data and analysis with support from quantum calculations and associated vibrational mode assignments, and supplementary references. See DOI: 10.1039/x0xx00000x

cell imaging. However, the structural origin of the emission redshift has remained elusive. We previously postulated that the formation of an RFP-like chromophore can expand the chromophore's π-conjugation via additional self-catalyzed oxidation,[14] while another study proposed that the environmental effect rearranges hydrogen (H)-bonding near the chromophore.[13] In this work, we aimed to gain structural insights and perform targeted engineering. We combined a series of spectroscopic, electrochemical, and computational techniques to investigate the chromophore structure and determined that a "double donor" mechanism which differs from previous propositions is mainly responsible for the unusually large redshift. Moreover, by using a more robust procedure for the aY-modified superfolder GFP (aY-sfGFP) expression, we found that the fluorescence quantum yield (FQY) of aY-sfGFP was overestimated.[14] Subsequently, we used the delineated mechanistic insights into the chromophore structure as well as protein engineering methods to significantly boost the FQY for aY-sfGFP. This effective feedback and discovery loop can benefit the crucial characterization and improvement of myriad photosensitive, complex systems from noncanonical RFPs to nanomachines across various disciplines.

2. Results and discussion

2.1. Origin of red emission in GFP-like model chromophores

We first compared the absorption and fluorescence properties of sfGFP and aY-sfGFP with the synthetic model chromophore of sfGFP, *para*-hydroxybenzylidene-dimethylimidazolinone (HBDI, Fig. 1A), and its *ortho*-amino-derivatized variant, amino-HBDI (Fig. 1B). We note that the two synthetic model chromophores have higher pK_a in water (e.g., ~8.4 for HBDI),[15] but when they (such as HBDI) are placed within FPs, the chromophore local environment is expected to reduce their pK_a values, resulting in deprotonation (reaching the anionic form) at neutral pH buffer conditions.[14,16] To make our comparisons more relevant, we specifically deprotonated the two synthetic model chromophores using alkaline solution before performing spectroscopic measurements. Notably, since the $-OH/-O^-$ pair (p$K_{a,1} \approx 8.1$) has a higher pK_a value than the $-NH_3^+/-NH_2$ pair (p$K_{a,2} \approx 4.3$) in amino-HBDI chromophore, under our condition it is unlikely to form the zwitterionic tautomer between $-OH$ and $-NH_2$, i.e., $-O^-/-NH_3^+$ in the ground state (see supporting data plots in Fig. S1, ESI†). Therefore, all the chromophores under discussion here (solution vs. FP environments) are the same anionic forms shown in Fig. 1.

From steady-state electronic spectroscopy, sfGFP and aY-sfGFP display similar but slightly lower emission peak energies than those of the model chromophores, HBDI and amino-HBDI, in acetonitrile (509 vs. 500 nm, and 605 vs. 575 nm), respectively (see Fig. 2A, top vs. bottom panels). Meanwhile, the absorption peaks of the two FPs show more pronounced redshifts with respect to the model chromophores (488 vs. 461 nm, and 541 vs. 499 nm). These experimental findings align with the fact that the protein local environment (electrostatic and steric effects) can impact on the electronic ground and excited states of the chromophore, as previously investigated via theoretical calculations.[17,18] In solution, the deprotonated amino-HBDI exhibits solvatochromism for both the electronic absorption and emission profiles, indicative of the excited-state intramolecular charge transfer (ICT) (see Fig. S2 and Table S1 in the ESI†). Kamlet-Taft solvatochromic analysis reveals that the spectral shift in absorption is governed by the solute-solvent H-bonding interactions while the emission is mainly contributed by both H-bonding and dipolar interactions (more prominent) between the chromophore and solvent (see Section S1.1, ESI†). Nevertheless, these spectral comparisons show that the single amino substitution can red-shift the GFP chromophore emission to a significant extent in solution.[19]

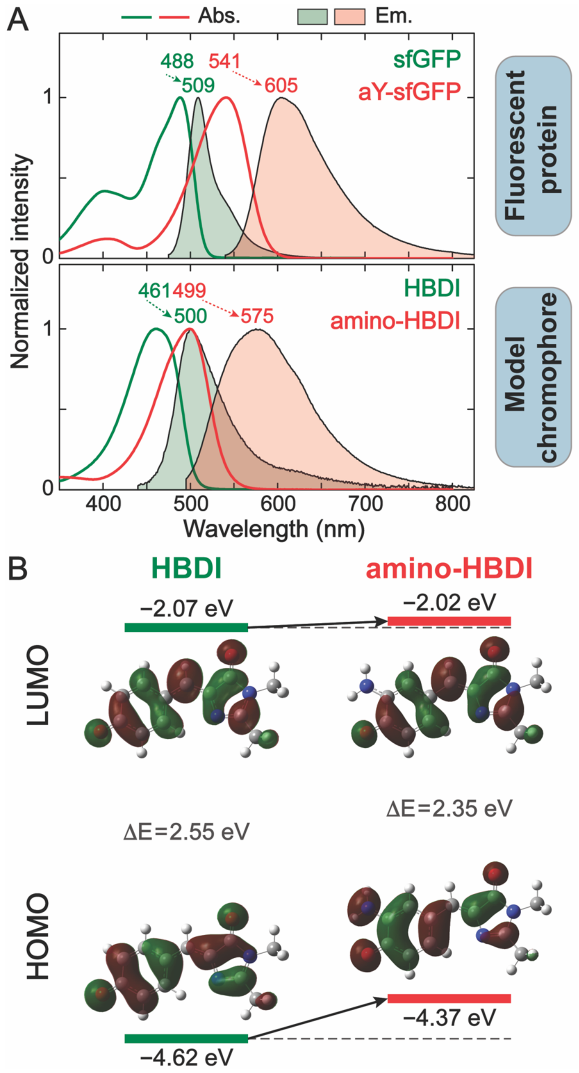

Fig. 2 Steady-state electronic spectra and energetics of sfGFP, aY-sfGFP, and their synthetic model chromophores. (A) Absorption and emission spectra of sfGFP/aY-sfGFP at pH 7.4 tris buffer with deprotonated chromophores and the deprotonated HBDI/amino-HBDI in acetonitrile with 0.1% (v/v) organic base, 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU). The spectra of model chromophores in acetonitrile are presented to be consistent with the solvent used in cyclic voltammetry. (B) Experimental energies measured by cyclic voltammetry (for the HOMO energy) as well as absorption/emission spectra (ΔE), and the computed electron density distribution of HOMO and LUMO for the deprotonated HBDI and amino-HBDI in acetonitrile with 0.1% (v/v) DBU.

The large redshift induced by a small substituent (mild structural change) is peculiar because the chromophore conjugation size is only slightly enlarged. To provide more experimental insights into the amino-caused change of electronic structure, we measured the HOMO energy of HBDI and amino-HBDI model chromophores in acetonitrile using the cyclic voltammetry method (see Fig. 2B and Fig. S3A,C, ESI†) and then estimated the LUMO energy by adding the 0-0 transition energy gap to the HOMO energy.[20,21] Moreover, the 0-0 transition energy gap (ΔE in Fig. 2B) was approximated experimentally as the intersection of the normalized reduced absorption and emission spectra (see Section S1.2 in the ESI† for detailed justification of this method, and Fig. S3B,D).

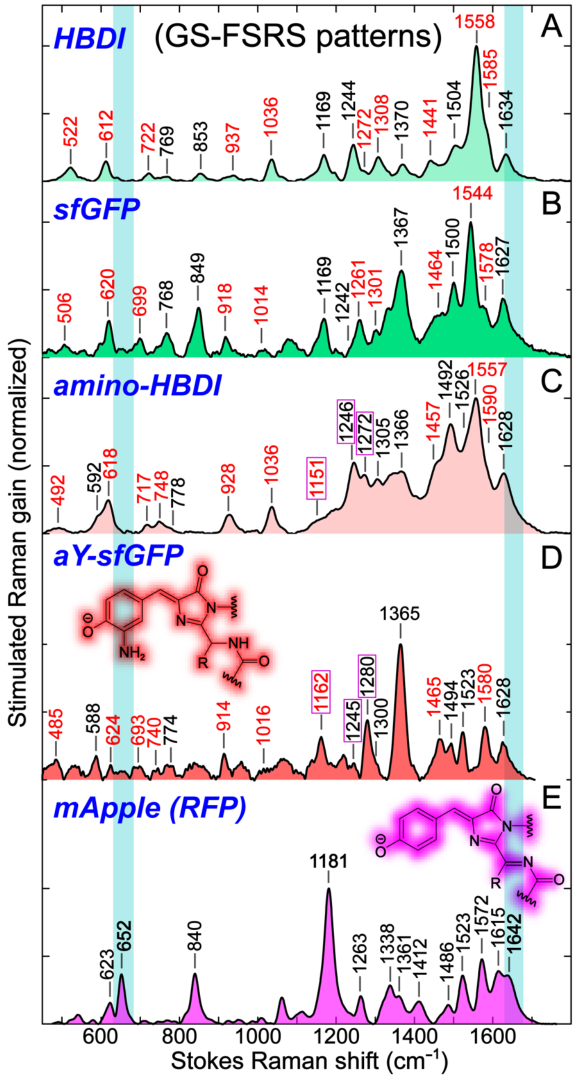

Fig. 3 Ground-state FSRS spectra of sfGFP, aY-sfGFP, mApple, and the synthetic model chromophores in solution. Peak frequencies are determined from least-squares fits using gaussian profiles. Shifted Raman mode frequencies due to $-CH_3$ motions are indicated in red. Several high-frequency modes with notable $-NH_2$ motions of the chromophore are highlighted by vertical magenta boxes (panels C/D). Cyan rectangular shades mark two characteristic RFP modes among the distinct Raman spectral patterns. All five spectra were recorded when the chromophores were deprotonated in pH 7.4 tris buffer for the FPs, and in aqueous solution with 1 mM NaOH for the model chromophores.

We stress that the 0-0 transition gap was used here as a qualitative estimate for the HOMO-LUMO gap that primarily contributes to the transition from the singlet ground (S_0) to first excited state (S_1) as substantiated by quantum calculations (Table S2, ESI†). The LUMO energy obtained this way reflects an averaged trend on the absorption and emission gaps since both optical transitions are not exactly given by a single HOMO-LUMO transition. The values and trend are corroborated by quantum calculations with the economical methods (ESI†, Fig. S4 and Table S2), which indicates that the large electronic peak redshift of amino-HBDI is governed by the drastically raised HOMO energy (–4.37 vs. –4.62 eV, see Fig. 2B). The energy increase is reflected by the altered electron density distribution showing more localized electron density mainly on the phenolate and $-NH_2$ group as well as an additional nodal plane between the two moieties. Notably, this mechanism is different from the redshift mechanism of RFPs containing an extended conjugation moiety at the imidazolinone end (e.g., N-acylimine), which causes the decreased excited-state energy mostly due to the enhanced phenolate-to-imidazolinone ICT.[19,22] Our results also align well with recent reports on color tuning of the GFP chromophore in solution and protein environments.[19,23] Meanwhile, the strong electron-donating capability of $-NH_2$ substituent in such a "double-donor" structure ($-O^-$ as the other donor, see Fig. 1B) causes a prominent ICT process upon the HOMO–LUMO transition which accounts for the pronounced solvatochromism of amino-HBDI (see below, and Section S1.1 in the ESI†). This inherent property may also explain the large Stokes shifts of amino-HBDI and aY-sfGFP (Fig. 2A), which make the latter protein even more advantageous in imaging applications besides the emission color.

2.2. Spectroscopic identification of the aY-sfGFP chromophore structure

Although the amino substitution alone demonstrates potential in dramatically red-shifting the GFP chromophore color in solution, we could not completely exclude the possibility of RFP-like chromophore formation as an add-on effect for an already red-shifted chromophore. This point requires particular caution when the model chromophore exhibits solvatochromism that the emission wavelength can significantly vary in different environments (Fig. S2, ESI†). Because of the resilience of aY-modified GFP to crystallization,[13] a direct identification of the chromophore structure is not readily available. Therefore, we implemented femtosecond stimulated Raman spectroscopy (FSRS)[24-26] with high structural sensitivity in the electronic ground state to systematically examine vibrational signatures for aY-sfGFP chromophore in contrasting local environments. The unique resonance enhancement in the tunable FSRS methodology enables us to obtain vibrational spectrum with high signal-to-noise ratio solely for the chromophore without contamination from other amino acids of the protein,[26] thus allowing for the unambiguous identification of the embedded chromophore structure by strategic comparisons with a series of control samples, particularly between the synthetic model chromophores in solution and their counterparts in FPs (Fig. 3).

Table 1 Representative high-frequency vibrational mode assignments for fluorescent proteins (FPs) and model chromophores

Prominent methyl motions? (Y or N)	$-NH_2$ motions involved? (Y or N)	HBDI	sfGFP	amino-HBDI	aY-sfGFP	Vibrational mode assignment [a]
N	Y	1634 [b]	1627 [b]	1628	1628	**P**($\nu_{C=C}$), **B**($\nu_{C=C}$, ρ_{C-H}), **I**($\nu_{C=O}$)
Y	Y	1585	1578	1590	1580	**I**($\nu_{C=N}$, δ_{C-H}, $\nu_{C=O}$), **P**(ρ_{C-H})
Y	N	1558	1544	1557	–	**P**($\nu_{C=C}$, $\nu_{C=O}$, ρ_{C-H}), **B**(ν_{C-C}, ρ_{C-H}), **I**($\nu_{C=N}$, $\nu_{C=O}$, δ_{C-H})
N	Y	1504	1500	1492	1494	**P**($\nu_{C=O}$, ρ_{C-H}), **B**($\nu_{C=C}$, ρ_{C-H}), **I**(δ_{C-H})
Y	Y	1441	1464	1457	1465	**I**(δ_{C-H}, ν_{C-N}), **P**(ρ_{C-H})
N	Y	1370	1367	1366	1365	**P**(ρ_{C-H}), **B**(ρ_{C-H}), **I**(ν_{C-N})

[a] Common vibrational motions between the deprotonated HBDI and amino-HBDI chromophores in water (i.e., the $-NH_2$ motions specific to amino-HBDI and aY-sfGFP are not shown). All the experimental Raman peak frequencies listed were measured from the ground-state FSRS spectra as shown in Fig. 3. The calculated full list of normal modes including the prominent methyl motions and characteristic $-NH_2$ motions in amino-HBDI can be seen in Tables S3 and S4 in the ESI†. Note that sfGFP and aY-sfGFP have the Thr-Tyr-Gly (TYG) chromophore (Fig. 1). The abbreviations in this table are for (1) structural moieties: **P** (phenolate ring), **I** (imidazolinone ring), **B** (methine bridge); and (2) characteristic vibrational motions: ν (stretching), δ (bending), ρ (in-plane rocking). [b] The 7 cm^{-1} difference experimentally observed between HBDI and sfGFP for this mode without prominent imidazolinone methyl motions (e.g., δ_{C-H} herein) might be due to other specific chromophore-environment interactions in sfGFP.

As the well-known model chromophore for GFP, HBDI exhibits similar mode frequencies to sfGFP due to prominent vibrations of the HBDI backbone (see peaks with black labels in Fig. 3A,B), except for the modes (red labels in Fig. 3A,B) that involve significant bending motions of the imidazolinone methyl groups (see example modes in Table 1 and the full Raman mode assignments in the ESI†, Fig. S5, and Tables S3 and S4). This useful pattern is because HBDI is truncated at these two methyl sites versus the sfGFP chromophore in the protein matrix (see Fig. 1A,C). Depending on the prominence of $-CH_3$ bending components in the normal mode, the observed peak frequency difference could vary (e.g., 23 cm^{-1} for 722/699 cm^{-1} vs. 7 cm^{-1} for 1308/1301 cm^{-1} in HBDI/sfGFP, see Fig. 3A/B and Table S3 in the ESI†). In analogy, by considering such characteristic frequency shifts (Table S4, ESI†), aY-sfGFP (Fig. 3D) exhibits similar mode frequencies to amino-HBDI (Fig. 3C), suggesting that the aY-sfGFP chromophore has the same structure as amino-HBDI (Fig. 1B,D).

In addition, we found that amino-HBDI spectrally resembles HBDI as the vibrational motions and frequencies of most Raman bands above 1000 cm^{-1} (mainly stretching modes of the HBDI backbone) are highly conserved between these two model chromophores (Fig. 3A and C, see mode assignments in the ESI†, Tables S3 and S4). This result is reminiscent of the halogenated GFP chromophores showing similar vibrational modes to GFP chromophore in the protein or solution environment,[15,27] which implies that a small substituent *ortho* to $-O^-$ on the phenolate ring does not induce a significant change in the ground-state vibrational motions of the HBDI backbone. To substantiate this useful point, we examined more substituents at this atomic site of a series of HBDI derivatives and found similar mode frequencies particularly in the high-frequency region, regardless of the electron-withdrawing or -donating substituent on the phenolate ring (see Fig. S6, ESI†). This finding supports the use of mApple with a different substituent on the imidazolinone ring (see below) as an effective control sample for the chromophore structure determination in aY-sfGFP. Moreover, the associated normal mode compositions of many Raman peaks change to various extents due to the involvement of $-NH_2$ motions (Fig. 3). For example, the 1441 cm^{-1} mode of HBDI shifts to 1457 cm^{-1} in amino-HBDI due to prominent in-plane bending motions of the $-NH_2$ group (Table 1). Such frequency shifts can also be seen in other modes (e.g., 1151, 1246, and 1272 cm^{-1}, highlighted by magenta boxes in Fig. 3C,D with detailed mode assignments listed in Tables S3 and S4, ESI†).

These correlated spectral results highlight the subtle yet detectable influence of a substituent at the donor moiety of GFP chromophore on the ground-state electronic and vibrational properties. Therefore, the detailed comparisons between HBDI (Fig. 3A) and amino-HBDI (Fig. 3C) need to focus on the Raman modes with conserved atomic motions (e.g., see Table 1), which display similar frequencies and are also shared by sfGFP (Fig. 3B) and aY-sfGFP (Fig. 3D) without and with the $-NH_2$ substituent, respectively.

As further experimental evidence, the DsRed-derivative mApple,[28] a typical RFP with the extended π-conjugation via N-acylimine double bonds ($-C=N-C=O$), exhibits globally different vibrational signatures from sfGFP and aY-sfGFP. Other RFPs with different extended conjugation structures at the imidazolinone end have also been reported to exhibit contrasting Raman spectra to GFP,[29,30] implying that the incorporation of an electron-withdrawing conjugated moiety to this imidazolinone ring site influences electron density distribution of the ground state and hence vibrational modes. This key point is further corroborated by substitution of an electron-withdrawing group at this site such as $–CF_3$ and $–COCH_3$ (i.e., the model chromophore of RFP, asFP595)[31,32] which leads to a notably distinct spectrum showing different normal modes and frequencies (see Fig. S7 in the ESI† for FSRS fingerprints), in contrast to the largely conserved modes across various phenolate-ring substituents (Fig. S6, ESI†). Therefore, even without the phenolate amino substituent, mApple (with its chromophore structure shown in Fig. 3E inset) still serves as a useful control sample to infer the chromophore structure in aY-sfGFP (Fig. 3D). The common RFP model chromophore with an extended acylimine or its phenolate *ortho*-amino derivative was not investigated here due to the substantial challenges encountered with unstable intermediates for its synthesis.

Specifically, in mApple, the prominent 652 cm^{-1} mode arises from an out-of-plane twist of the acylimine C=N bond while the 1642 cm^{-1} mode results from the acylimine C=N/C=O stretch coupled with imidazolinone C=O stretch, both of which are characteristic of a "conventional" RFP chromophore with the acylimine moiety (Fig. S8B, ESI†).[5,33] Furthermore, other marker bands involving the acylimine motions include ~1615, 1486, and 1181 cm^{-1}, just to name a few (Fig. 3E, see mode assignments in Fig. S8 and Table S5 in the ESI†). In essence, no clear observation of these marker bands across a wide spectral window for the pre-resonantly enhanced Raman peaks in FSRS (Fig. 3D, highlighted by semi-transparent cyan shades) strongly suggests the absence of an N-acylimine structure in aY-sfGFP.

We note that other RFP-like chromophore structures such as the asFP595 chromophore with a conjugated acetyl group (–C=O)[31,32] should also be unlikely for aY-sfGFP in view of the remarkable impact of a conjugated electron-withdrawing substituent at the imidazolinone ring end on the chromophore vibrational modes as shown (Fig. S7 in the ESI†, see vertical gray lines for example) and in literature.[30] One remaining point is that the mode intensity pattern of aY-sfGFP seems to differ from amino-HBDI especially in the high-frequency region above 1000 cm^{-1} (Fig. 3C,D). However, this observable is less indicative than mode frequency because the apparent Raman mode intensities of the same chromophore can vary with the resonance condition, vibronic effect, and local environment.[25,34] For instance, a change in the FP local environment alters the Raman mode intensity pattern to a significant extent as observed in two single mutants of aY-sfGFP in this work (see below, and more details in the ESI†). Nevertheless, systematic and robust comparisons of the FSRS marker band frequencies (an intrinsic property)[25,35] of a series of related FPs and their model chromophores across a wide spectral range (see Fig. 3, and Figs. S5–S8 in the ESI†) provide compelling evidence for a GFP-like instead of an RFP-like chromophore inside aY-sfGFP.

Therefore, our new results provide fundamental physical chemistry insights into an emergent ncAA-RFP, and correct the previous proposition that the redshift of aY-sfGFP with respect to sfGFP stems from the formation of an RFP chromophore via self-catalyzed oxidation.[14] The other mechanism suggesting the decisive role of H-bonding in the local chromophore environment[13] is very unlikely considering the massive magnitude of redshift (absorption: ~490 nm in sfGFP vs. 540 nm in aY-sfGFP; emission: 510 nm in sfGFP vs. 605 nm in aY-sfGFP) as presented in Fig. 2A top panel. Moreover, the conjugation length of GFP chromophore poses an intrinsic limit to its spectral wavelengths in different environments. The reddest canonical FP with an HBDI chromophore structure reported to date is phiYFP (525 nm in absorption and 537 nm in emission)[36,37] according to the FP database.[38] The absorption peak's red limit at ~525 nm has been accurately predicted in Boxer group's recent work that quantitatively modeled the relationships between the electro-optical properties of GFP and chromophore environment.[23] In our work herein, we also mutated aY-sfGFP to change the H-bonding network with or near the chromophore which only leads to ca. 10–30 nm spectral shifts (see below). Therefore, the change of protein chromophore structure should dominate the large redshift for aY-sfGFP as delineated by the aforementioned FSRS signatures and systematic comparisons with robust analysis, while the local environment variation plays a minor role.

2.3. Devising generalizable strategies to improve protein brightness

Upon optimizing conditions for preparing the relatively large quantity of aY-sfGFP needed for FSRS study, we noticed that aY-sfGFP could be readily contaminated with sfGFP. In fact, aY is a redox-sensitive amino acid and its undesired oxidation by molecular oxygen in air could cause varied results. Using sealed culture containers, a 1:5 medium-to-container volume ratio, and induction at a high bacterial optical density resulted in aY-sfGFP with minimal sfGFP contaminant (see Experimental Section below). We re-examined photophysical properties of the freshly purified aY-sfGFP and found that the brightness of aY-sfGFP (FQY=0.037) was previously overestimated.[14] Since the low brightness poses a drawback for imaging applications, we next sought in this work to engineer aY-sfGFP to boost its brightness.[4,39,40] To dissect the chromophore-environment interactions and rationally improve aY-sfGFP, we first studied amino-HBDI that is essentially nonfluorescent in aqueous solution (FQY≈10^{-4}) due to the isomerization-induced rapid nonradiative decay pathways.[15,41,42] Interestingly, amino-HBDI displays fluorogenic behavior in solution that smaller solvent polarity causes higher FQY (Table S1, ESI†). This result indicates prominent ICT, so the isomerization barrier of fluorescent state (FS) toward transition state (TS) is sensitive to environmental polarity. The negative correlation between FQY and polarity suggests a more polar structure of TS than FS in amino-HBDI, so higher solvent polarity leads to more populations moving out of FS toward an S_1/S_0 conical intersection.[43,44] A closer inspection suggests that the steric effect may also contribute to FQY. The correlation between FQY (ϕ) and electronic (polarity, E_T^N) and steric (viscosity, η) effects can be modeled by a linear relationship: $\log(1/\phi - 1) \sim A \cdot E_T^N + B \cdot \log(\eta)$ (see Section S1.3, ESI†).[45] The best-fit results ($A > 0$, $B < 0$) manifest that smaller polarity and larger viscosity of the solvent increase the chromophore FQY. This finding can be visualized by the single-parameter plot as we define an electro-steric parameter: $\sigma =$

$E_T^N + B/A \cdot \log(\eta)$ so $\log(1/\phi - 1) \sim A \cdot \sigma$ (see Fig. 4A, and Table S6 in the ESI†), and the linear relationship is conspicuous for amino-HBDI in various solvents.

Caution is needed when one translates knowledge in a chromophore FQY between different microenvironments. In solution, FS and TS experience relatively homogeneous electrostatic fields applied by solvent molecules,[45] resulting in a decreased FS-to-TS barrier in polar solvents for amino-HBDI. Protein environment around the chromophore, however, is inhomogeneous[26,46,47] and could exert distinct impacts on FS and TS, and alter the FQY differently. Notably, TS adopts a twisted chromophore structure along the isomerization coordinate for these HBDI derivatives.[26,44] With these essential mechanistic insights, we devise feasible strategies to stabilize FS (strategy 1) or destabilize TS (strategy 2) of the protein chromophore via electrostatic or steric confinements as illustrated in Fig. 4B, which is nontrivial since the potential energy surface is affected by different factors at different locations as the photoexcited chromophore searches phase space to reach the fluorescent state.

The logical primary targets for mutagenesis involve residues proximal to the chromophore due to their strong interactions. In particular, the symmetry breaking of a substituted phenolate results in two possible conformations with $-NH_2$ on site 3 or 5 referred to as conformation 1 and 2, respectively (Fig. 1D). We note that the ground-state FSRS data cannot readily discriminate between the two conformers as our current level of calculations yield similar spectra (Fig. S9, ESI†). However, quantum mechanics/molecular mechanics (QM/MM) calculations and molecular dynamics (MD) simulations have supported conformation 2 for aY-modified avGFP (a structural analogue of sfGFP) due to the H-bonding network with nearby polar residues like S205.[13] Considering the similar chromophore pockets for these two proteins, aY-sfGFP likely also adopts conformation 2 (see Fig. S10B, ESI†).[48] The fact that our strategic mutations at this particular side of the chromophore can improve FQY (see below) further substantiates the assignment. We next sought for specific protein residues for mutagenesis to examine the aforementioned two strategies (Fig. 4B) that are generalizable to other FPs to improve FQY and advance their bioimaging potentials.

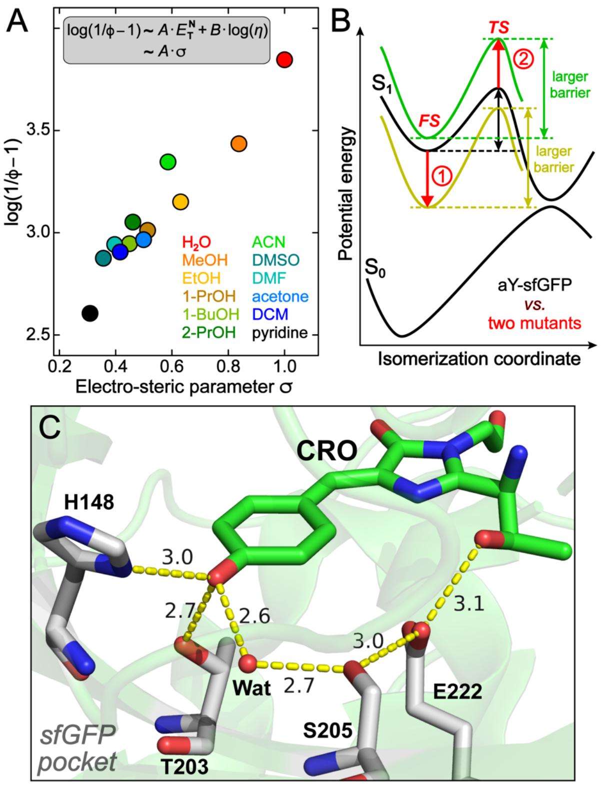

Fig. 4 Correlations between the chromophore FQY and environmental factors. (A) Dependence of amino-HBDI's FQY on solvent properties. (B) Schematic energy diagram proposed for tuning the FQY. The double-sided arrows depict the chromophore isomerization barriers. Red arrows denote the energy tuning for FS and TS, separately, reaching the illustrative potential energy surfaces (brown and green solid curves) for strategies 1 and 2. (C) Local environment of the sfGFP chromophore (PDB ID: 2B3P;[48] see more details in Fig. S10, ESI†). The relevant H-bonds are indicated by yellow dashed lines with distances marked in Å unit.

2.4. Rational mutagenesis of aY-sfGFP for brighter mutants

The deprotonated chromophores of GFPs usually exhibit high FQYs, mainly attributed to the H-bonding wire that connects the chromophore Y66, water, S205, E222, and S65/T65 (Fig. 4C).[49-51] The weakened H-bonding like in avGFP-E222Q decreases the chromophore FQY because glutamine is a weaker H-bond acceptor.[50] This result supports our model that the H-bonding wire stabilizes FS. The drastic FQY drop from sfGFP (~0.65) to aY-sfGFP (0.037) insinuates that the H-bonding wire is disrupted in aY-sfGFP, wherein the $-NH_2$ group on the phenolate ring approaches S205 and E222 (Fig. S10B in the ESI†) to force a reorganization of the H-bonding network near the chromophore. In particular, $-NH_2$ likely forms an H-bond with S205 and breaks the one between S205 and E222 (glutamate in pH 7.4 buffer) that can only act as an H-bond acceptor. This scenario is reminiscent of the aY-modified avGFP wherein E222 significantly deviates from its original position in avGFP.[13] We hypothesized that rebuilding an H-bond between site 222 and S205 or even $-NH_2$ could stabilize FS and increase FQY (strategy 1, Fig. 4B). Among the isosteric residues,[52] histidine represents a suitable candidate because of its H-bond-donating capability. Through site-directed mutagenesis, the resultant single mutant E222H exhibits the remarkably improved brightness with FQY=0.425, comparable to other bright RFPs.[4,40,53] The blue-shifted absorption and emission peaks (at ~522 and 575 nm, see Table 2 and Fig. S11 in the ESI†) corroborate the strengthened H-bonding with the chromophore that effectively stabilizes the negative charge of phenolate as supported by recent work on the color-tuning mechanism for GFP,[23,49] while the red emission still characterizes the mutant as a potent RFP. Further evidence arises from the ground-state FSRS studies on various proteins,[26] wherein the Raman mode frequency and intensity show correlated patterns with the H-bonding and π–π stacking interactions (see Fig. S12 and Section S1.4 in the ESI†). These important results provide detailed evidence for our above-mentioned interpretation that a change of FP local environment (even with a single point mutation) could exert drastic effects on the vibrational mode intensity (see notable Raman peak

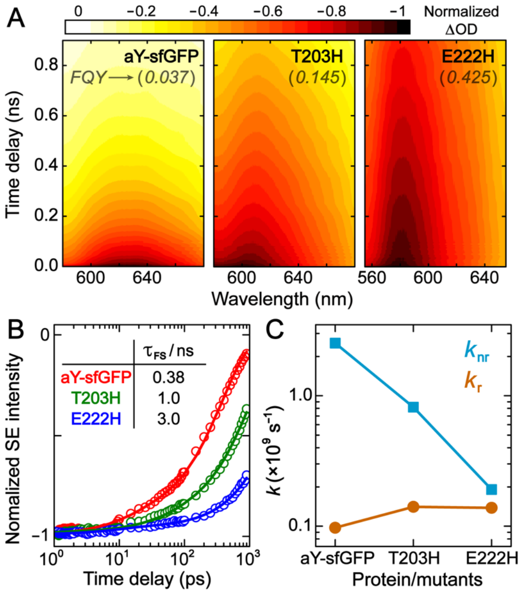

Fig. 5 Excited-state decay of aY-sfGFP and its brighter mutants. (A) Contour plots of femtosecond transient absorption spectra in the stimulated emission (SE) region. (B) Intensity dynamics of the protein SE band with decay time constants listed for three proteins in the inset. (C) Rate constants of the nonradiative (k_{nr}, cyan) and radiative (k_r, orange) decay pathways retrieved in three proteins show distinct trends.

pattern changes in Fig. S12A, ESI†), thus making it less indicative or useful than the Raman mode frequency in revealing the chromophore structure as manifested in Fig. 3.[25,26]

Alternatively, a destabilized TS can increase the isomerization barrier (see design strategy 2 in Fig. 4B). Due to the relevant twisted structure, we reason that placing a bulky residue below or above the chromophore conjugation plane can cause steric clash and thus destabilize TS. Potential candidates include the common histidine or tyrosine with aromatic rings at site 203 to interact with the chromophore phenolate ring. The pertinent π–π stacking interactions have been found in various FPs with high FQYs.[39,54-57] The resultant T203H mutant shows an improved FQY of 0.145 (see Table 2, and Fig. S11 in the ESI†), but T203Y is dimmer than aY-sfGFP. It implies that T203Y does not favor strong π–π stacking, likely due to the symmetry-breaking $-NH_2$ substituent on the chromophore phenolate ring with a bulky Y203 nearby.

The increased FQYs of these newly engineered protein mutants correlate with their lengthened excited-state lifetimes (k_{FS}) from ultrafast transient absorption spectroscopy (see Fig. 5A,B, and Fig. S13 in the ESI† for global analysis results). Calculations of the nonradiative (k_{nr}) and radiative (k_r) decay rate constants (i.e., $k_{nr} = (1-\phi)k_{FS}$ and $k_r = \phi k_{FS}$, see Section S1.3 in the ESI†) reveal that the improved FQY mainly originates from the decrease in k_{nr} with a largely unchanged k_r (Fig. 5C), which validates our rational design strategies for improving FQY via effective inhibition of the chromophore isomerization that constitutes the major nonradiative pathway. Last but not least, we expressed the double mutant T203H/E222H to check whether an additive effect exists for the FQY. Interestingly, T203H/E222H shows a similar FQY (0.415) to E222H (Table 2, and Fig. S11 in the ESI†) but a drastically decreased extinction coefficient (Fig. S14, ESI†), making its overall brightness inferior to E222H. The very low extinction coefficient of aY-sfGFP-T203H/E222H may indicate a distorted chromophore (non-planar) in the ground state[58,59] due to the two strongly interacting bulky histidine residues in close proximity within the protein matrix. This property can be detrimental to the FQY since no additive effect on FQY was observed, which could inspire future investigations.

Table 2 Photophysical properties of aY-sfGFP and amino-HBDI

	λ_{abs} / nm	λ_{em} / nm	ϕ (FQY)
amino-HBDI [a]	499	575	4.5×10^{-4}
aY-sfGFP [b]	541	605	0.037
T203H [b]	540	601	0.145
E222H [b]	522	575	**0.425**
T203H/E222H [b]	535	580	0.415

[a] Measured in acetonitrile solvent with 0.1% (v/v) DBU (see Table S1 in the ESI†). [b] Measured in pH 7.4 tris buffer (see Fig. S11 in the ESI†).

3. Materials and methods

3.1. Plasmid construction, mutagenesis, and protein expression

The noncanonical amino acid 3-aminotyrosine (*H*-3-amino-Tyr-OH·2HCl, Cat # 402789) was purchased from Bachem (Torrance, CA, USA).[14] All oligonucleotides were synthesized by Integrated DNA Technologies (Coralville, Iowa, USA). Sanger sequencing was performed by Eurofins Genomics (Louisville, KY, USA) for plasmid confirmation.

To perform mutagenesis at residue positions 205 and 222, pBAD-sfGFP(Y66TAG) was separately amplified with Arab_f and S205NNK_r, or Arab_r and E222NNK_f. The two fragments purified from gel extraction were used for three-fragment Gibson assembly[60] along with a pBAD plasmid (Thermo Fisher) predigested with Hind III and Xho I.

To perform mutagenesis at residue position 203, pBAD-sfGFP(Y66TAG) was separately amplified with Arab_f and T203NNK_r, or Arab_r and T203NNK_f. The two fragments purified from gel extraction were used for three-fragment Gibson assembly along with a pBAD plasmid (Thermo Fisher) predigested with Hind III and Xho I.

The gene libraries were used to transform E. cloni 10G competent cells (Lucigen) containing the pEvol-MjaYRS plasmid (Addgene, Cat # 153557) via electroporation. Cells were plated on 2×YT agar plates supplemented with 100 µg/mL ampicillin, 35 µg/mL chloramphenicol, 4 mM aY amino acid, and 0.2% (w/v) L-arabinose. After incubation at 37 °C overnight, the plates were sealed with parafilm and incubated at room temperature for 48–72 h. A laboratory-built colony fluorescence imaging system (excitation filter 550/30 nm, emission filter 645/75 nm) was used to image the red fluorescence of the colonies. Bright colonies were selected and used to inoculate 300 µL of Terrific Broth (TB) medium supplemented with 100 µg/mL ampicillin and 35 µg/mL chloramphenicol in a 96-deep well plate sealed with adhesive foil (VWR, Cat # 60941-126) at 37 °C overnight. Next, another 100 µL of TB containing 0.8% (w/v) L-arabinose and 16 mM aY amino acid was added, resulting in ~400 µL culture containing 0.2% (w/v) L-arabinose and 4 mM aY in every 2 mL well. The plate was sealed again with adhesive foil to minimize amino acid oxidation before being cultured at 30 °C and 220 rpm for another 48 h.

Cells were pelleted by centrifugation, and the pellets were lysed with Bacterial Protein Extraction Reagents (BPER, Thermo Fisher) on ice with gentle shaking for 30 min. The supernatants were used for screening on a BioTek Synergy Mx Microplate Reader. Fluorescence (excitation 540 nm, emission 605 nm) was measured, and promising colonies were sequenced. pBAD-sfGFP(Y66TAG)-T203H and pBAD-sfGFP(Y66TAG)-E222H were generated during the library screening processes that corroborate our proposed rational design strategies (see main text above) for enhancing the brightness of the protein fluorescence. To evaluate the potentially additive effect, the double-mutant (pBAD-sfGFP(Y66TAG)-T203H/E222H) was generated from pBAD-sfGFP(Y66TAG)-E222H using the oligonucleotides Arab_f, T203H_r, Arab_r, and T203H_f.

The sequences (5′→3′) of the oligonucleotides mentioned above are:

Arab_f: CTTGACGGCTACATCATTC
S205NNK_r:
CACGTTTTTCATTCGGATCTTTGCTCAGAACMNNCTGGGTGCTCAGATA
Arab_r: GAATGATGTAGCCGTCAAG
E222NNK_f:
GATCCGAATGAAAAACGTGATCATATGGTGCTGCTGNNKTTTGTTACCGCCGCG
T203NNK_r:
GCTCAGAACGCTCTGMNNGCTCAGATAATGATTATC
T203NNK_f:
GATAATCATTATCTGAGCNNKCAGAGCGTTCTGAGC
T203H_r: GCTCAGAACGCTCTGGTGGCTCAGATAATGATTATC
T203H_f: GATAATCATTATCTGAGCCACCAGAGCGTTCTGAGC

To express the relatively large quantity of proteins for all the spectroscopic measurements, the pBAD plasmids were used to transform E. cloni 10G competent cells (Lucigen) containing the pEvol-MjaYRS plasmid. A single colony of each mutant was picked and grown in a capped culture tube with 1.0 mL TB supplemented with 100 µg/mL ampicillin and 35 µg/mL chloramphenicol at 37 °C and 220 rpm overnight. The saturated starter was then diluted 100-fold into 100 mL TB supplemented with 100 µg/mL ampicillin and 35 µg/mL chloramphenicol in a 500 mL flask. For optimal protein expression, the medium volume should be equal to ~1/5 of the volume of the container. When the optical density (OD) at 600 nm reached 2 per cm, 0.2% L-arabinose and 4 mM aY amino acid were added to the culture. The flask was sealed and incubated at 30 °C and 220 rpm for 48 h. Cells were pelleted, resuspended in 1×PBS (phosphate-buffered saline, pH 7.4), and lysed by sonication. His-tagged proteins were first enriched and purified with Ni-NTA agarose beads (Pierce, Rockford, IL), then subjected to an Akta protein purification system (Cytiva) that was equipped with a HiLoad 16/600 Superdex 200 pg size-exclusion column. Proteins were kept in the elution buffer (30 mM Tris HCl, 150 mM NaCl, pH 7.4) at 4 °C, and further analysis was completed within 1–3 weeks.

3.2. Spectroscopic measurements and analysis

The steady-state electronic absorption spectra of all the samples were collected using a Thermo Scientific Evolution 201 UV/Visible spectrophotometer. The fluorescence spectra were collected using a Shimadzu RF-6000 spectrofluorophotometer. The solution sample measured was filled in a four-sided, 10-mm-pathlength quartz cuvette. All the spectroscopic measurements were conducted at room temperature (22 °C) and standard pressure (1 atm).

The femtosecond transient absorption (fs-TA) and ground-state femtosecond stimulated Raman spectra (FSRS) for all samples were collected using a home-built optical setup on the basis of a mode-locked Ti:sapphire oscillator and a laser regenerative amplifier (Legend Elite-USP-1K-HE; Coherent, Inc.), which provides an ~800 nm fundamental pulse (FDP) with 35 fs duration and 1 kHz repetition rate. The detailed description of both spectroscopic setups can be found in our previous publications.[61-63] In brief, visible fs pump pulses in fs-TA were generated through a two-stage fs noncollinear optical parametric amplifier (NOPA), followed by temporal compression via a chirped mirror pair (DCM-12, 400−700 nm, Laser Quantum, Inc.).[64] The fs probe from supercontinuum white light (SCWL) was generated by focusing a small portion of FDP onto a 2-mm-pathlength quartz cuvette filled with deionized water, and temporally compressed by a chirped mirror pair (DCM-9, 450−950 nm, Laser Quantum, Inc.). The cross-correlation time between the fs actinic pump and SCWL probe was ~100 fs. The sample was housed in a 1-mm-pathlength quartz cuvette and flowed through soft tubing that was pumped by a mini electric motor. The pump wavelength was strategically tuned to 490 nm for sfGFP, 510 nm for E222H, 520 nm for T203H/E222H, and 535 nm for aY-sfGFP and T203H, in accord with their absorption profiles (Fig. S11A, ESI†). The pump power used was ~0.2–0.3 µJ/pulse for all samples. Global analysis of the fs-TA spectra was performed using the open-access Glotaran program.[65] A sequential kinetic model that yields the evolution-associated difference spectrum (EADS)[66,67] was used to retrieve intrinsic rate constants underlying the chromophore excited-state electronic dynamics (Fig. S13, ESI†).

The ground-state FSRS experiment consists of two laser pulses: a picosecond (ps) Raman pump that was generated by a two-stage ps-NOPA system, and an fs probe (SCWL) that was shared with the aforementioned fs-TA setup but compressed by a chirp mirror pair of either DCM-12 (400−700 nm, Laser Quantum, Inc.) for sfGFP, HBDI, and amino-HBDI samples or DCM-9 (450−950 nm, Laser Quantum, Inc.) for aY-sfGFP, T203H, E222H, T203H/E222H, and mApple samples (i.e., the samples with redder absorption and emission profiles in this work). To achieve pre-resonance enhancement to increase the vibrational signal-to-noise ratio for robust spectral analysis, the Raman pump wavelength was tuned to 534 nm (HBDI, amino-HBDI, sfGFP), 580 nm (E222H), 589 nm (T203H/E222H), 601 nm (aY-sfGFP), 610 nm (T203H), and 615 nm (mApple). For the two series of synthetic model chromophores, the singly substituted HBDI chromophores at the phenolate (P)-ring (Fig. S6, ESI†) or imidazolinone (I)-ring (Fig. S7, ESI†),[15,19,32,68] the Raman pump wavelengths were selected accordingly to achieve pre-resonance condition and with decent Raman pump average power (3–4 mW) to obtain high signal-to-noise ratio. The pre-resonance Raman pump wavelength at the very red edge of the electronic absorption band is expected to generate minimal contribution from the excited-state vibrations. This point is corroborated by our previous systematic investigations of the ground-state FSRS line shapes using the Rhodamine 6G dye.[69] The excited-state contributions (e.g., peak frequency shift and width broadening) can only be observed with Raman pump wavelengths that significantly go into the absorption band. All the spectra were collected on the Stokes side of the visible Raman pump, and processed consistently with spectral baseline subtraction to obtain Raman peaks as shown in Fig. 3.[25,26] The sample solutions were all housed in a 1-mm-pathlength quartz cuvette (Spectrosil 1-Q-1; Starna Cells, Inc.).

3.3. Cyclic voltammetry

Cyclic voltammetry measurements were performed with a CHI bipotentiostat. Gold (2 mm in diameter) and stainless-steel wires were used as the working and counter electrodes, respectively, along with the silver (Ag) wire as the quasi-reference electrode. The samples (1 mM HBDI and amino-HBDI) were prepared by dissolving the compound in dry acetonitrile and adding 0.1% (v/v) DBU (1,8-diazabicyclo[5.4.0]undec-7-ene) to achieve the deprotonation of the chromophore in solution. Tetrabutylammonium hexafluorophosphate ($[Bu_4N]PF_6$, 0.1 M) was added as the supporting electrolyte. Ferrocenemethanol (FcMeOH) was used as the reference. The cyclic voltammograms were obtained at a scan rate of 50 mV/s to obtain the onset potentials for the deprotonated HBDI and amino-HBDI in acetonitrile (ESI†, Section S1.2 and Fig. S3A,C) as the model chromophores.

3.4. Computational methods

To qualitatively predict the energy shift trend (not focusing on the exact values) for electronic transitions of the model chromophore due to a single amino substituent, ground- and excited-state calculations for the anionic (i.e., deprotonated) HBDI and amino-HBDI were performed with density functional theory (DFT) and time-dependent (TD)-DFT at the B3LYP/6-311G+(d,p) functional level/basis set, respectively. The chromophore geometry was optimized for both the electronic ground (S_0) and first excited (S_1) singlet states with the implicit integral equation formalism variant polarizable continuum model (IEFPCM) to account for solvation effects. Water and acetonitrile were used as IEFPCM solvents. All calculations were performed using the Gaussian 16 program.[70] The S_0–S_1 absorption and emission are dominantly contributed by the HOMO–LUMO transition for these molecules (Table S2 and Fig. S4 in the ESI†), and the associated electron density distribution changes can provide useful insights into the observed electronic absorption and emission peak shifts between HBDI derivatives (see Fig. 2 for example). The calculated ground-state Raman spectra were frequency-scaled to match the experimental FSRS spectra across a wide spectral window for a series of comparative samples from the model chromophores to their corresponding chromophores in the protein matrix (see Figs. S5 and S8, and Tables S3–S5 in the ESI†).

4. Conclusions

In this contribution, we implemented the wavelength-tunable FSRS methodology in the electronic ground state, aided by other complementary spectroscopic and computational techniques, to characterize an emergent noncanonical RFP and reveal a GFP-like chromophore inside aY-sfGFP. Our findings substantiate the dominant role of chromophore structural change over environmental effect on the large redshift of aY-sfGFP. The electron-donating amino substitution of the GFP chromophore alone is adequate to greatly red-shift the fluorescence via a unique double-donor structure that raises the ground-state electronic energy and enhances intramolecular charge transfer after photoexcitation. In addition, we demonstrate that the chromophore environment has significant impacts on the fluorescence efficiency. To improve the aY-sfGFP brightness, we devised rational design and engineering strategies with fundamental electronic and steric insights from a unified excited-state potential energy scheme. The ~12-fold increase in FQY of aY-sfGFP-E222H makes it feasible to advance bioimaging such as FRET biosensing.[3,53,56] We envision that implementing the effective substitution and mutation strategy based on deep molecular insights in this work to existing RFPs may further shift their emission wavelengths to far-red or near-IR regions with high brightness. We also anticipate that high-level calculations such as all-atom or ab initio MD simulations and hybrid QM/MM methods can provide deeper insights into the amino- and other noncanonical chromophore structure and its interactions with a dynamic local environment,[18,71-74] further advancing rational protein design for myriad applications in bioimaging and life sciences.

Data availability

The datasets with analysis and discussions supporting this article have been incorporated and presented as part of the comprehensive ESI†.

Author contributions

Conceptualization, C. F. and H.-w. A.; methodology, C. C., H. Z. and J. Z.; software, C. C. and C. F.; formal analysis, C. C. and H. Z.; investigation, C. C., H. Z. and J. Z.; visualization, C. C., H. Z. and C. F.; validation, C. C., H. Z., J. Z., H.-w. A. and C. F.; writing–original draft, C. C. and H. Z.; writing–review and editing, C. F. and H.-w. A.; supervision, C. F.; funding acquisition, C. F. and H.-w. A. All authors have read and agreed to the published version of the manuscript.

Conflicts of interest

There are no conflicts of interest to declare.

Acknowledgements

We thank the financial support by U.S. National Science Foundation (NSF) grants MCB-1817949 and CHE-2003550 (to C. F.), and U.S. National Institutes of Health (NIH) grants R01GM129291, U01CA230817, and R01DK122253 (to H.-w. A.). We appreciate Dr. Mikhail Baranov for providing the synthetic model chromophores and other GFP chromophore derivatives, and Dr. Partha Sheet for helping with the CV measurements.

Notes and references

1 M. V. Matz, A. F. Fradkov, Y. A. Labas, A. P. Savitsky, A. G. Zaraisky, M. L. Markelov and S. A. Lukyanov, *Nat. Biotechnol.*, 1999, **17**, 969-973.

2 G. S. Baird, D. A. Zacharias and R. Y. Tsien, *Proc. Natl. Acad. Sci. U. S. A.*, 2000, **97**, 11984-11989.

3 R. E. Campbell, O. Tour, A. E. Palmer, P. A. Steinbach, G. S. Baird, D. A. Zacharias and R. Y. Tsien, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 7877-7882.

4 N. C. Shaner, R. E. Campbell, P. A. Steinbach, B. N. G. Giepmans, A. E. Palmer and R. Y. Tsien, *Nat. Biotechnol.*, 2004, **22**, 1567-1572.

5 F. V. Subach and V. V. Verkhusha, *Chem. Rev.*, 2012, **112**, 4308-4327.

6 A. S. Mishin, F. V. Subach, I. V. Yampolsky, W. King, K. A. Lukyanov and V. V. Verkhusha, *Biochemistry*, 2008, **47**, 4666-4673.

7 A. M. Bogdanov, A. S. Mishin, I. V. Yampolsky, V. V. Belousov, D. M. Chudakov, F. V. Subach, V. V. Verkhusha, S. Lukyanov and K. A. Lukyanov, *Nat. Chem. Biol.*, 2009, **5**, 459-461.

8 A. Sattarzadeh, R. Saberianfar, W. R. Zipfel, R. Menassa and M. R. Hanson, *Sci. Rep.*, 2015, **5**, 11771.

9 J. H. Bae, M. Rubini, G. Jung, G. Wiegand, M. H. J. Seifert, M. K. Azim, J.-S. Kim, A. Zumbusch, T. A. Holak, L. Moroder, R. Huber and N. Budisa, *J. Mol. Biol.*, 2003, **328**, 1071-1081.

10 L. Wang, X. Chen, X. Guo, J. Li, Q. Liu, F. Kang, X. Wang, C. Hu, H. Liu, W. Gong, W. Zhuang, X. Liu and J. Wang, *Biophy. Rep.*, 2018, **4**, 273-285.

11 J. K. Villa, H.-A. Tran, M. Vipani, S. Gianturco, K. Bhasin, B. L. Russell, E. J. Harbron and D. D. Young, *Molecules*, 2017, **22**, 1194.

12 M. G. Romei, C.-Y. Lin, I. I. Mathews and S. G. Boxer, *Science*, 2020, **367**, 76-79.

13 G. Augustine, S. Raghavan, K. NumbiRamudu, S. Easwaramoorthi, G. Shanmugam, J. Seetharani Murugaiyan, K. Gunasekaran, C. Govind, V. Karunakaran and N. Ayyadurai, *J. Phys. Chem. B*, 2019, **123**, 2316-2324.

14 S. Zhang and H.-w. Ai, *Nat. Chem. Biol.*, 2020, **16**, 1434-1439.

15 S. A. Boulanger, C. Chen, L. Tang, L. Zhu, N. S. Baleeva, I. N. Myasnyanko, M. S. Baranov and C. Fang, *Phys. Chem. Chem. Phys.*, 2021, **23**, 14636-14648.

16 L. M. Oltrogge, Q. Wang and S. G. Boxer, *Biochemistry*, 2014, **53**, 5947-5957.

17 V. R. I. Kaila, R. Send and D. Sundholm, *Phys. Chem. Chem. Phys.*, 2013, **15**, 4491-4495.

18 A. Petrone, P. Caruso, S. Tenuta and N. Rega, *Phys. Chem. Chem. Phys.*, 2013, **15**, 20536-20544.

19 C. Chen, M. S. Baranov, L. Zhu, N. S. Baleeva, A. Y. Smirnov, S. O. Zaitseva, I. V. Yampolsky, K. M. Solntsev and C. Fang, *Chem. Commun.*, 2019, **55**, 2537-2540.

20 Y. Li, Y. Cao, J. Gao, D. Wang, G. Yu and A. J. Heeger, *Synth. Met.*, 1999, **99**, 243-248.

21 H. Li, J. Li, D. Liu, T. Huang and D. Li, *Chem. Eur. J.*, 2020, **26**, 6899-6909.

22 Z. Wang, Y. Zhang, C. Chen, R. Zhu, J. Jiang, T.-C. Weng, Q. Ji, Y. Huang, C. Fang and W. Liu, *Angew. Chem. Int. Ed.*, 2023, **62**, e202212209.

23 C.-Y. Lin, M. G. Romei, L. M. Oltrogge, I. I. Mathews and S. G. Boxer, *J. Am. Chem. Soc.*, 2019, **141**, 15250-15265.

24 C. Fang, R. R. Frontiera, R. Tran and R. A. Mathies, *Nature*, 2009, **462**, 200-204.

25 C. Fang, L. Tang, B. G. Oscar and C. Chen, *J. Phys. Chem. Lett.*, 2018, **9**, 3253–3263.

26 C. Fang and L. Tang, *Annu. Rev. Phys. Chem.*, 2020, **71**, 239-265.

27 B. G. Oscar, L. Zhu, H. Wolfendeen, N. D. Rozanov, A. Chang, K. T. Stout, J. W. Sandwisch, J. J. Porter, R. A. Mehl and C. Fang, *Front. Mol. Biosci.*, 2020, **7**, 131.

28 N. C. Shaner, M. Z. Lin, M. R. McKeown, P. A. Steinbach, K. L. Hazelwood, M. W. Davidson and R. Y. Tsien, *Nat. Methods*, 2008, **5**, 545-551.

29 S. G. Kruglik, V. Subramaniam, J. Greve and C. Otto, *J. Am. Chem. Soc.*, 2002, **124**, 10992-10993.

30 D. C. Loos, S. Habuchi, C. Flors, J.-i. Hotta, J. Wiedenmann, G. U. Nienhaus and J. Hofkens, *J. Am. Chem. Soc.*, 2006, **128**, 6270-6271.

31 M. L. Quillin, D. M. Anstrom, X. Shu, S. O'Leary, K. Kallio, D. M. Chudakov and S. J. Remington, *Biochemistry*, 2005, **44**, 5774-5787.

32 I. V. Yampolsky, S. J. Remington, V. I. Martynov, V. K. Potapov, S. Lukyanov and K. A. Lukyanov, *Biochemistry*, 2005, **44**, 5788-5793.

33 K. Nienhaus, H. Nar, R. Heilker, J. Wiedenmann and G. U. Nienhaus, *J. Am. Chem. Soc.*, 2008, **130**, 12578-12579.

34 A. B. Myers and R. A. Mathies, Resonance Raman Intensities: A Probe of Excited-State Structure and Dynamics, in *Biological Applications of Raman Spectroscopy*, ed. T. G. Spiro, John Wiley & Sons, Inc., New York, 1987, vol. 2, pp. 1-58.

35 D. R. Dietze and R. A. Mathies, *ChemPhysChem*, 2016, **17**, 1224–1251.

36 D. A. Shagin, E. V. Barsova, Y. G. Yanushevich, A. F. Fradkov, K. A. Lukyanov, Y. A. Labas, T. N. Semenova, J. A. Ugalde, A. Meyers, J. M. Nunez, E. A. Widder, S. A. Lukyanov and M. V. Matz, *Mol. Biol. Evol.*, 2004, **21**, 841-850.

37 C. Chen, J. N. Tutol, L. Tang, L. Zhu, W. S. Y. Ong, S. Dodani and C. Fang, *Chem. Sci.*, 2021, **12**, 11382–11393.

38 T. J. Lambert, *Nat. Methods*, 2019, **16**, 277-278.

39 E. M. Merzlyak, J. Goedhart, D. Shcherbo, M. E. Bulina, A. S. Shcheglov, A. F. Fradkov, A. Gaintzeva, K. A. Lukyanov, S. Lukyanov and T. W. J. Gadella, *Nat. Methods*, 2007, **4**, 555-557.

40 D. S. Bindels, L. Haarbosch, L. van Weeren, M. Postma, K. E. Wiese, M. Mastop, S. Aumonier, G. Gotthard, A. Royant, M. A. Hink and T. W. J. Gadella, Jr. , *Nat. Methods*, 2017, **14**, 53-56.

41 N. M. Webber, K. L. Litvinenko and S. R. Meech, *J. Phys. Chem. B*, 2001, **105**, 8036-8039.

42 M. A. Taylor, L. Zhu, N. D. Rozanov, K. T. Stout, C. Chen and C. Fang, *Phys. Chem. Chem. Phys.*, 2019, **21**, 9728-9739.

43 C. Chen, S. A. Boulanger, A. I. Sokolov, M. S. Baranov and C. Fang, *Chemosensors*, 2021, **9**, 234.

44 C. Chen and C. Fang, *Chemosensors*, 2023, **11**, 87.

45 C. Chen, S. R. Tachibana, N. S. Baleeva, I. N. Myasnyanko, A. M. Bogdanov, A. S. Gavrikov, A. S. Mishin, K. K. Malyshevskaya, M. S. Baranov and C. Fang, *Chem. Eur. J.*, 2021, **27**, 8946-8950.

46 T. D. Krueger, L. Tang, L. Zhu, I. L. Breen, R. M. Wachter and C. Fang, *Angew. Chem. Int. Ed.*, 2020, **59**, 1644-1652.

47 M. Drobizhev, R. S. Molina, P. R. Callis, J. N. Scott, G. G. Lambert, A. Salih, N. C. Shaner and T. E. Hughes, *Front. Mol. Biosci.*, 2021, **8**, 633217.

48 J.-D. Pédelacq, S. Cabantous, T. Tran, T. C. Terwilliger and G. S. Waldo, *Nat. Biotechnol.*, 2006, **24**, 79-88.

49 K. Brejc, T. K. Sixma, P. A. Kitts, S. R. Kain, R. Y. Tsien, M. Ormö and S. J. Remington, *Proc. Natl. Acad. Sci. U. S. A.*, 1997, **94**, 2306-2311.

50 G. Jung, J. Wiehler and A. Zumbusch, *Biophys. J.*, 2005, **88**, 1932-1947.

51 A. Royant and M. Noirclerc-Savoye, *J. Struct. Biol.*, 2011, **174**, 385-390.

52 D. Auerbach, M. Klein, S. Franz, Y. Carius, C. R. D. Lancaster and G. Jung, *ChemBioChem*, 2014, **15**, 1404-1408.

53 G. C. H. Mo, C. Posner, E. A. Rodriguez, T. Sun and J. Zhang, *Nat. Commun.*, 2020, **11**, 1848.

54 R. M. Wachter, D. Yarbrough, K. Kallio and S. J. Remington, *J. Mol. Biol.*, 2000, **301**, 157-171.

55 H.-w. Ai, S. G. Olenych, P. Wong, M. W. Davidson and R. E. Campbell, *BMC Biol.*, 2008, **6**, 13.

56 A. J. Lam, F. St-Pierre, Y. Gong, J. D. Marshall, P. J. Cranfill, M. A. Baird, M. R. McKeown, J. Wiedenmann, M. W. Davidson, M. J. Schnitzer, R. Y. Tsien and M. Z. Lin, *Nat. Methods*, 2012, **9**, 1005-1012.

57 J. Akerboom, N. Carreras Calderón, L. Tian, S. Wabnig, M. Prigge, J. Tolö, A. Gordus, M. B. Orger, K. E. Severi, J. J. Macklin, R. Patel, S. R. Pulver, T. J. Wardill, E. Fischer, C. Schuler, T.-W. Chen, K. S. Sarkisyan, J. S. Marvin, C. I. Bargmann, D. S. Kim, S. Kugler, L. Lagnado, P. Hegemann, A. Gottschalk, E. R. Schreiter and L. L. Looger, *Front. Mol. Neurosci.*, 2013, **6**, 2.

58 F. V. Subach, L. Zhang, T. W. J. Gadella, N. G. Gurskaya, K. A. Lukyanov and V. V. Verkhusha, *Chem. Biol.*, 2010, **17**, 745-755.

59 S. Pletnev, F. V. Subach, Z. Dauter, A. Wlodawer and V. V. Verkhusha, *J. Mol. Biol.*, 2012, **417**, 144-151.

60 D. G. Gibson, L. Young, R.-Y. Chuang, J. C. Venter, C. A. Hutchison and H. O. Smith, *Nat. Methods*, 2009, **6**, 343-345.

61 W. Liu, F. Han, C. Smith and C. Fang, *J. Phys. Chem. B*, 2012, **116**, 10535-10550.

62 L. Zhu, W. Liu and C. Fang, *Appl. Phys. Lett.*, 2014, **105**, 041106.

63 W. Liu, Y. Wang, L. Tang, B. G. Oscar, L. Zhu and C. Fang, *Chem. Sci.*, 2016, **7**, 5484-5494.

64 L. Tang, L. Zhu, M. A. Taylor, Y. Wang, S. J. Remington and C. Fang, *Molecules*, 2018, **23**, 2226.

65 J. J. Snellenburg, S. P. Laptenok, R. Seger, K. M. Mullen and I. H. M. van Stokkum, *J. Stat. Softw.*, 2012, **49**, 1-22.

66 I. H. M. van Stokkum, D. S. Larsen and R. van Grondelle, *Biochim. Biophys. Acta*, 2004, **1657**, 82-104.

67 C. Chen, L. Zhu, M. S. Baranov, L. Tang, N. S. Baleeva, A. Y. Smirnov, I. V. Yampolsky, K. M. Solntsev and C. Fang, *J. Phys. Chem. B*, 2019, **123**, 3804-3821.

68 M. S. Baranov, K. M. Solntsev, K. A. Lukyanov and I. V. Yampolsky, *Chem. Commun.*, 2013, **49**, 5778-5780.

69 C. Chen, L. Zhu and C. Fang, *Chin. J. Chem. Phys.*, 2018, **31**, 492-502.

70 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, Williams, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman and D. J. Fox, *Gaussian 16, Rev. C.01*, Gaussian, Inc., Wallingford, CT, 2016.

71 K. B. Bravaya, B. L. Grigorenko, A. V. Nemukhin and A. I. Krylov, *Acc. Chem. Res.*, 2012, **45**, 265-275.

72 B. L. Grigorenko, A. V. Nemukhin, I. V. Polyakov, D. I. Morozov and A. I. Krylov, *J. Am. Chem. Soc.*, 2013, **135**, 11541-11549.

73 G. Donati, A. Petrone, P. Caruso and N. Rega, *Chem. Sci.*, 2018, **9**, 1126-1135.

74 F. Coppola, F. Perrella, A. Petrone, G. Donati and N. Rega, *Front. Mol. Biosci.*, 2020, **7**, 569990.