# Interplay between Locally Excited and Charge Transfer States Governs the Photoswitching Mechanism in Fluorescent Protein Dreiklang

Tirthendu Sen[a], Yingying Ma[a], Igor V. Polyakov[b,c], Bella L. Grigorenko[b,c], Alexander V. Nemukhin[b,c], Anna I. Krylov[a]

[a] Department of Chemistry, University of Southern California, Los Angeles, California 90089-0482

[b] Department of Chemistry, Lomonosov Moscow State University, Moscow, Russia

[c] Emanuel Institute of Biochemical Physics, Russian Academy of Sciences, Moscow, Russia

We present the results of high-level electronic structure and dynamics simulations of the photoactive protein Dreiklang. With the goal to understand the details of Dreiklang's photocycle, we carefully characterize the excited states of the ON- and OFF-forms of Dreiklang. The key finding of our study is the existence of a low-lying excited state of a charge-transfer (CT) character in the neutral ON form and that population of this state, which is nearly isoenergetic with the locally excited (LE) bright state, initiates series of steps ultimately leading to the formation of the hydrated dark chromophore (OFF state). These results allow us to refine the mechanistic picture of the Dreiklang's photocycle and photoactivation.

## I. INTRODUCTION

Many fluorescent proteins (FPs) undergo reversible photoswitching upon photoexcitation, which is instrumental in several imaging modalities[1], including super-resolution techniques[2,3]. The most common mechanism is cis-trans photoisomerization of the chromophore, sometimes coupled with changes in its protonation state; notable examples include Dronpa[4], Padron[5], and KFP[6]. An entirely different mechanism is operating in Dreiklang[7], where the switching is based on reversible photoinduced hydration/dehydration of the imidazolinone ring of the chromophore (Fig. 1).

Dreiklang[7] was derived by random mutagenesis from Citrine, a close relative of EYFP[8]. It has the same chromophore formed by the glycine-tyrosine-glycine (GYG) tripeptyde $\pi$-stacked with a nearby tyrosine residue (Tyr203). The chromophore's conjugated core (Fig.

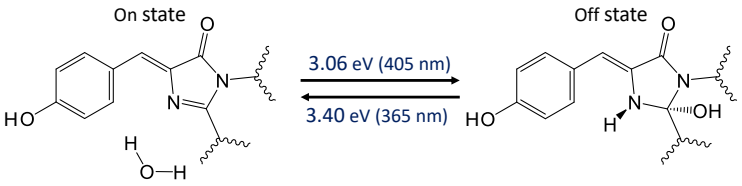


FIG. 1: On-off photoconversion in Dreiklang is activated by photoexcitation of the neutral form of the chromophore in ON-state. The OFF-form can be turned on by photoexcitation at higher energy.

1) is the same as in EGFP[9], but due to the T65G mutation the connection to peptide backbone via imidazolinone's carbon is slightly different.

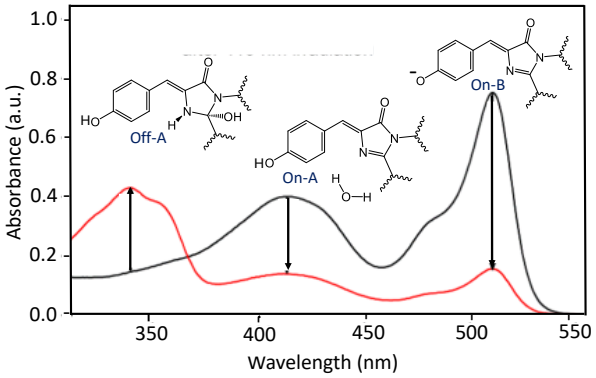


FIG. 2: Steady-state absorption spectra of the ON-state (black) and following irradiation (red) at 3.02 eV (410 nm) at pH 7.5. The spectra are from Ref. 10. The band maxima are at 3.01 eV and 2.43 eV in the ON-state and at 3.65 eV in the OFF-state.

Fig. 2 shows absorption spectra of the ON- and OFF- forms of Dreiklang[7,10]. The absorption spectrum of the ON-form features two bands: one at $\sim$ 3.01 eV (411-413 nm) and a twice more intense one at 2.43 eV (511 nm), with a shoulder at 2.58 eV (480 nm). These bands are assigned to the neutral and anionic forms of the chromophore, traditionally called form A and form B[11,12]. In many other GFP-like proteins[11], excitation to either band leads to the identical fluorescence spectra with the maximum at around 2.3 eV (green or yellow), ascribed to form B. This is explained by ultrafast (picoseconds or shorter) excited-state proton-transfer (ESPT) from the chromophore to a proton acceptor via a proton wire[11,13,14]. In wt-GFP, ESPT proceeds as a sequential proton transfer from the excited neutral chromophore to the Glu222 carboxylate through a water molecule and the hydroxyl group of

Ser205[12,15–18]. This pathway can be disrupted by targeted mutations.

In Dreiklang, excitation of peak B leads to fluorescence at 2.34 eV (529 nm), with the quantum yield of 0.41. However, in contrast to many other FPs, the excitation of peak A leads to very weak (albeit non-negligible[10]) fluorescence. This weak steady-state fluorescence is identical[10] to the fluorescence produced by excitation of peak B. These observations suggest that ESPT in Dreiklang is strongly suppressed and happens with a very small quantum yield. Reduced effectiveness of ESPT is consistent with the observation[7] that the essential difference between the parent system (Citrine) and Dreiklang is the upshift of the pKa of the ON-state of the chromophore (7.2 versus 5.7).

The distinguishing feature of Dreiklang is that irradiation of peak A results in photoconversion to the dark form (OFF-state)[7]. Thus, in imaging applications, the fluorescence of Dreiklang can be excited by using 2.43 eV light, recorded at 2.34 eV, and turned off by 3.01 eV light. This decoupling of the fluorescence excitation from photoswitching makes Dreiklang very attractive and is responsible for its name (Dreiklang is the German word for a three-note chord in music). The OFF-state spontaneously returns to the ON-state in the course of minutes. Alternatively, the dark state can be switched back by irradiation at 3.65 eV (340 nm).

The X-ray structures[7] of the ON- and OFF-states (PDB IDs: 3ST2/3ST4 and 3ST3, respectively) show that the ON-state is indeed similar to EGFP/EYFP, whereas the OFF-state has a hydrated chromophore, similar to an intermediate form of the immature chromophore[19] but with the methyne double bond.

The exact details of the photoswitching mechanism remain unknown. Espagne and coworkers[10] investigated the mechanism using transient absorption and concluded that formation of photoproducts occurs on a nanosecond time scale or slower. They reported spectroscopic evidence of the formation of excited-state (on a ps timescale) and ground-state (picosecond to nanosecond timescale) intermediates and proposed a tentative mechanism; however, the proposed structures of the intermediates have not been validated by theoretical modeling of their spectral properties. On the theoretical side, we investigated[20] thermal (ground-state) recovery of the ON-state. The calculations predicted the reaction barrier of about 27 kcal/mol and identified Glu222 as the key residue involved in the recovery reaction, while the scan of the excited-state surface suggested a barrierless OFF→ON photoreaction. This work also presented a cursory analysis of the structures of the ON- and OFF-states,

including tentative assignment of the protonation states of the key residues around the chromophore, and computed their spectral properties.

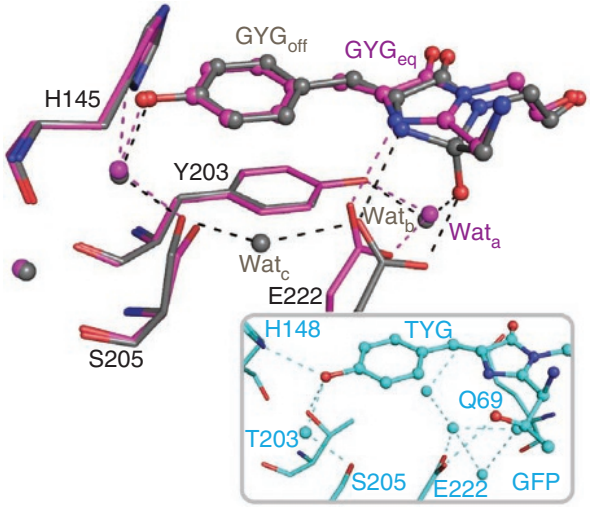


FIG. 3: Superimposed representations of the hydrogen-bond network around Dreiklang's chromophore in the ON- and OFF-states. Color scheme: ON-state carbons, magenta; OFF-state carbons, gray; oxygen, red; nitrogen, blue. Important water molecules are shown as magenta (ON-state) and gray (OFF-state) spheres. Inset: hydrogen-bond network in EGFP. *Reproduced from Ref. 7.*

Given the importance of proton wires in the photocycle of FPs, here we revisit the question of protonation states using more advanced computational protocols and assess the effect of different protonation states on the excited states of the chromophore. Dreiklang operates in a wide range of pH (6-9). Figure 3 shows superimposed crystal structures[7] of the ON-state (equilibrium structure, PDB ID 3ST2) and OFF-state (PDB ID 3ST3), indicating the hydrogen-bond network around the chromophore. It also compares the network around the Dreiklang chromophore with that in EGFP (PDB ID 1EMA)[9]. Dreiklang's structure clearly shows participation of His145, Glu222, and Ser205, as well as several water molecules. In contrast, in EGFP chromophore forms hydrogen bonds with His148 (position 145 is occupied by Tyr145 in EGFP), Thr203, and Glu222.

The two critical residues near the chromophore binding site in Dreiklang are Glu222 and His145. A tentative mechanism proposed in Ref. 10 assumed that both Glu222 and His145 are in the neutral form, at least, in the ON-state. In our study of the thermal recovery reaction[20], we considered Glu222 to be deprotonated and pointed out that the change in

its protonation state along the reaction profile plays an essential role. Given the significant differences in hydrogen-bond network in Dreiklang relative to EGFP, protonation states of the key residues should be carefully re-evaluated. We use the structure as the primary gauge and compare the distances between the selected residues and the chromophore. In some cases (e.g., for structures that have exactly the same atoms in the QM part), we also consider total energies of the optimized structures.

After obtaining model structures for different forms of the chromophore and for different protonation states of His145 and Glu222, we computed excitation energies and analyzed the effect of the protein environment. The key finding is that in the neutral form of the chromophore there is low-lying state of charge-transfer (CT) character (Tyr203→Chro), corresponding to electron transfer from Tyr203 to the chromophore. This state is only present in the neutral form. We further investigate an implications of this states by dynamical simulations and geometry optimizations. Our results indicate that population of the CT state, which is located within 0.25 eV from the bright locally excited (LE) state of $\pi\pi^*$ character, plays the key role in Dreklang's photoswitching. On the basis of our calculations, we propose a refined picture of photoconversion mechanism, summarized in Fig. 4. As discussed below, this mechanism is consistent with all available experimental findings[7,10].

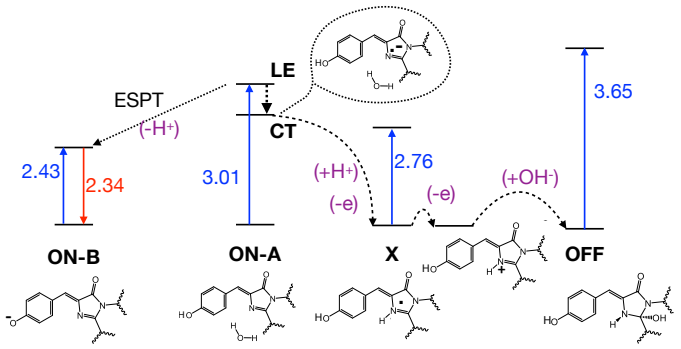


FIG. 4: Revised Dreiklang's photocycle. Excitation of form A can lead to ESPT and fluorescence, but this channel is suppressed in Dreiklang. Alternatively, the locally excited chromophore can undergo a non-adiabatic transition to the CT state, which is then stabilized by proton transfer. After releasing the electron back to Tyr203, intermediate X undergoes nucleophilic attack by nearby water, forming hydrated chromophore.

The structure of the paper is as follows. We begin by describing computational protocols. We then discuss the results of the simulations using different computational protocols. We

first consider different protonation states and the excited states of the chromophore. We then discuss excited-state dynamics of the chromophore and the role of the CT state in the photoconversion. We conclude by discussing the implications of the revised photocycle.

## II. COMPUTATIONAL METHODS AND PROTOCOLS

We begin with the crystal structure of the recovered ON-state (3ST4), which is nearly exactly superimposable on the equilibrium ON-structure (3ST2)[7]. The structure includes two water molecules: W354 near the phenolate end and W242 near the imidazolinone moiety. We note that in the previous study[20] we used 3ST2 as the starting point for the ON-state and the model structure also included additional water molecule, which is present in 3ST3 structure (OFF-state) but not seen in 3ST2 and 3ST4.

We consider the following protonation states: Chromophore is anionic or neutral in the ON-state and is neutral in the OFF-state. Depending on the local environment, His145 can have 3 different protonation states: HSD (protonated at $N_\delta$), HSE (protonated at $N_\epsilon$), and HSP (protonated on both N, positively charged). Glu222 can be GLU (anionic) or GLUP (protonated). Figures S1 and S2 in the SI summarize the names and definitions of different protonation states. *Propka*[21] suggested HSP state (pKa 2.2) for His145 and GLUP state (pKa 9.2) for the Glu222, however, as described below, our calculations indicated that alternative protonation states are more likely.

We build the model structures as follows. Starting from the PDB structure, hydrogen atoms were added using the *VMD* plugin and a modified (to include the chromophore) *CHARMM27* topology file. Protonation states were initially assigned by *Propka*[21] and then manually set for the chromophore, His145, and Glu222. Charged amino acids on the surface were locally neutralized by adding counterions close ($\sim$4.5 Å) to them. Charged residues that do not form salt bridges inside the protein barrel were also neutralized by adding appropriate counter ions at the surface. For HSD-GLUP structures, this protocol resulted in the addition of 19 $Na^+$ and 12 $Cl^-$ in the neutral forms (ON- and OFF-states), and 19 $Na^+$ and 11 $Cl^-$ in the anionic forms. For other protonation states the number of counter ions was adjusted accordingly. The proteins were solvated in water boxes producing a buffer of 15 Å. TIP3P water model was used to describe water. Molecular dynamics (MD) simulations were performed using these solvated neutralized model structures as follows:

1. Minimization using steepest descent algorithm for 2000 steps (protein, crystal water, counterions).

2. Minimization using steepest descent algorithm for 2000 steps of the fully solvated structure (keeping protein frozen), with the subsequent equilibration of the solvent (keeping the protein frozen) for 500 ps with 1 fs time step using NPT ensemble.

3. Full equilibration of the system for 2 ns (with 1 fs time step) with periodic boundary condition (PBC) using the isobaric-isothermal NPT ensemble (Noose-Hoover barostat with Langevin dynamics).

4. Production run for 2 ns with 1 fs time step using the NPT ensemble. Pressure and temperature were kept at 1 atm and 298 K.

The structures from production-run MD simulations were used to compute average structural parameters. We also used 21 snapshots from MD simulations to compute QM/MM excitation energies; in these calculations, the geometry of the QM part was not optimized.

To obtain better structures for more accurate estimate of the excitation energies, we carried out QM/MM optimizations using ONIOM, starting from the final structures from Step 1. To reduce the system size, in these calculations we removed the counterions and pruned the solvation shell, only retaining waters within 4 Å from the surface of the protein. The total size of the system in these calculations was ∼5900 atoms and the charge -7 (for the neutral ON form in HSE-GLUP state).

In the MD and QM/MM simulations we used *CHARMM27* parameters for standard protein residues[22] and the parameters derived by Reuter *et al.* for the anionic GFP chromophore[23]. The parameters for the hydrated form of the chromophore were derived from additional quantum mechanical calculations (optimized structures and NBO charges), as described in the SI. QM/MM optimizations were carried out using the mechanical embedding scheme (ONIOM). The definitions of the QM part used in ONIOM is shown in Fig. 5 (large QM). All coordinates were allowed to relax, except for the positions of link atoms ($C_\alpha$ carbons of the amino-acid residues shown in Fig. 5), which were pinned to the positions from the MM-relaxed structures.

The QM part was described by ωB97X-D/aug-cc-pVDZ in the QM/MM optimizations and in the AIMD (ab initio MD) simulations. This functional[24,25] belongs to the family

of long-range corrected functionals in which the notorious self-interaction error (SIE) is greatly reduced[26–29]; it also includes dispersion correction[30]. The benchmarks illustrated excellent performance of $\omega$B97x-D for structures and energy differences of a broad range of compounds[24,25]. Using LRC functionals is particularly important for charged systems and for describing CT states.

Excitation energies were computed using a finite cluster approach with slightly larger QM system (extended QM, see Fig. 5). Excitation energies were computed using several electronic structure methods: TD-DFT with $\omega$B97X-D, SOS-CIS(D)[31], EOM-CCSD[32], and XMCQDPT2[33]. In these calculations we used the following basis sets: cc-pVDZ, aug-cc-pVDZ on all atoms, and a mixed basis set, aug-cc-pVDZ on the heavy atoms of the chromophore and Tyr203 and cc-pVDZ on the rest of the atoms. The charge of the the large QM and extended QM is +1 for the neutral forms of the chromophore (due to the positively charged arginine) and zero for the anionic forms for all protonation states of the His148 and Glu222 except HSD-GLU, HSE-GLU, and HSP-GLUP (see Fig. S2 in the SI for the definition of protonation states).

The XMCQDPT2 perturbation theory calculations were based on the CASSCF wave functions obtained by distributing 16 electrons over 12 orbitals and using density averaging.The active space included orbitals from the chromophore and Tyr203.

In addition, we computed excitation energies using electrostatic embedding, as in our previous studies[34–37]. To prevent the overpolarization of the QM part, the charges on the boundary atoms were redistributed as follows[36,37]: bonds before -CONH were cut and capped with hydrogen atoms and the charge on CONH was set to zero; the excess charge was then redistributed on other atoms of the residue to maintain the total charge of the amino acid. These calculations were done using 21 snapshots from the MD trajectories (step 4 above) and the large QM system with the aug-cc-pVDZ basis set.

Fig. 5 shows the QM parts used in the ONIOM optimizations (large QM) and in the calculations of excitation energies (large QM and extended QM). We also carried calculations with minimal QM (chromophore), and with the medium QM (chromophore and Tyr203).

Excited-state AIMD simulations were performed using the same protocol as geometry optimization (ONIOM embedding, large QM, $\omega$B97X-D/aug-cc-pVDZ, CHARMM27 forcefield), with constant energy ensemble and using initial velocities corresponding to 298 K thermal distribution with 1 fs time step for 10 ps (10,000 steps).

All electronic structure calculations were carried out with Q-Chem[38,39], except for XMC-QDPT2 calculations, which were carried out with *Firefly*[40]. MD simulations were performed with NAMD[41].The excited-state analysis was carried out using the *libwfa* library[42]. In the SI, we also present the results for the structures from Ref. 20, which were obtained with a different QM/MM protocol.

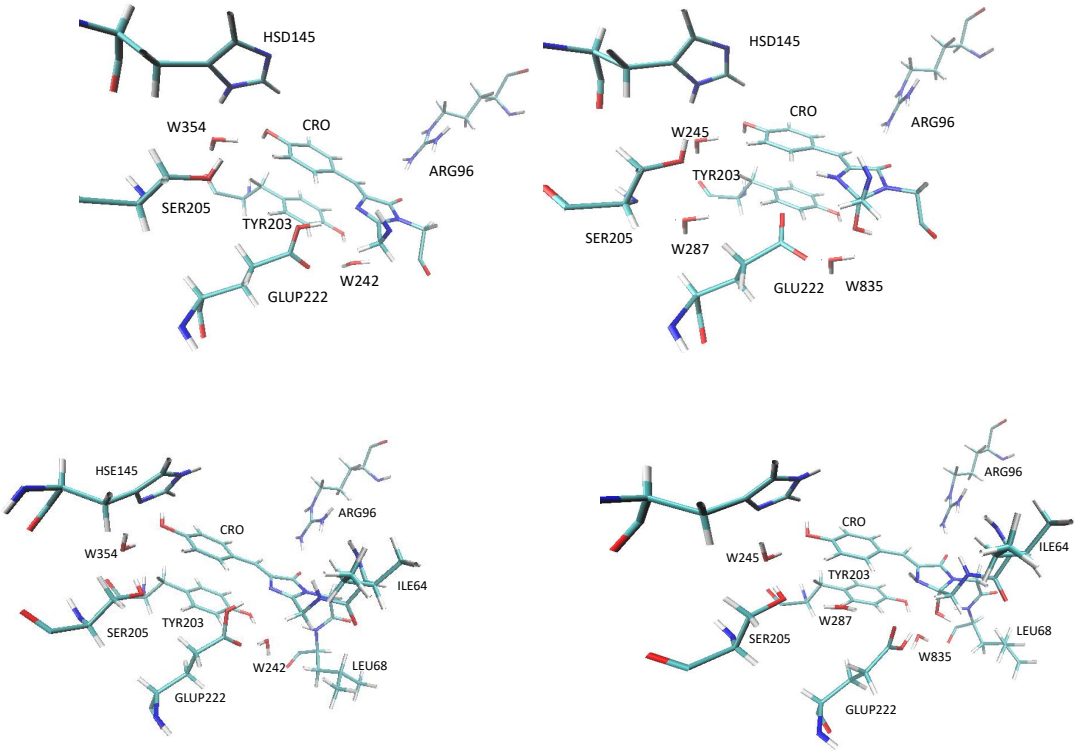


FIG. 5: Top: Residues included in large QM in the QM/MM calculations of ON- (left) and OFF-states. Bottom: Residues included in extended QM in the excited-state calculations of ON- (left) and OFF-states. Small QM contains only the chromophore and medium QM contains the chromophore and Tyr203. The total charge of the small and medium QM is zero for the A-form (neutral chromophore) and -1 for the B-form (anionic chromophore). For large and extended QM, the total charge of the QM is +1 for the A-form (ON and OFF) and zero for the B-form for all protonation states of the His148 and Glu222 except HSD-GLU, HSE-GLU, and HSP-GLUP (see Fig. S2 in the SI for the definition of protonation states). The numbering corresponds to the 3ST4 (left) and 3ST2 (right) crystal structures. For on-A (HSE-GLUP) structure, large and extended QM comprised 113 and 118 atoms, respectively.

## III. RESULTS AND DISCUSSION

### A. Protonation states for the ON-state

It is instructive to begin by revisiting the hydrogen-bonding network around EGFP chromophore. As clearly seen in Fig. 3, the EGFP network comprises Glu222, Ser205, Thr203, and His148. In EGFP, position 145 is occupied by tyrosine (not shown in the figure), which does not form a hydrogen bond with the chromophore. The protonation states of the key residues in EGFP (in the anionic form) are well established: Glu222 is protonated (neutral) and His148 is neutral (HSD form, protonated at $N_\delta$).[12,36,37]. In the neutral form, Glu222 is deprotonated[12] and the protonation state of His148 is the same as in the anionic form. We note that alternative protonation states are thermodynamically accessible and can be populated, especially at different pH.

In Dreiklang, Thr203 is replaced by tyrosine, which participates in $\pi$-stacking instead of hydrogen bonding. This difference has a major effect on the distance between Glu222 and Ser205: compare 4.18 Å in Dreiklang and 3.72 Å in EGFP. Another important difference is that His145 forms a hydrogen bond with the chromophore and His145 moves away. Furthermore, T65G substitution, which, as was shown recently[37], significantly weakens the hydrogen-bonding network around the chromophore, increasing its flexibility in the excited state.

Figures 6 and 7 show the definition of the key distances used to validate the structures of the ON- and OFF-states. Tables S7-S12 in the SI contain the average values computed along the MD trajectories, the values at the QM/MM optimized structures, and compare them with the respective values from the crystal structures. These values are presented graphically in Figs. 8 and 9. Fig. S6 in the SI shows relative energies of the optimized structures for the model systems where the QM parts contains the same set of atoms, such that the total energies are comparable.

We note that the comparison with crystal structure is complicated by the equilibrium between the anionic and neutral chromophores. The averaged distances from the MD simulations generally agree well with the values from QM/MM optimization, which provides validation of the force-field parameters; the largest differences are observed for d7 and d11 (water position). For the ON-state with the neutral chromophore, we observe the best agree-

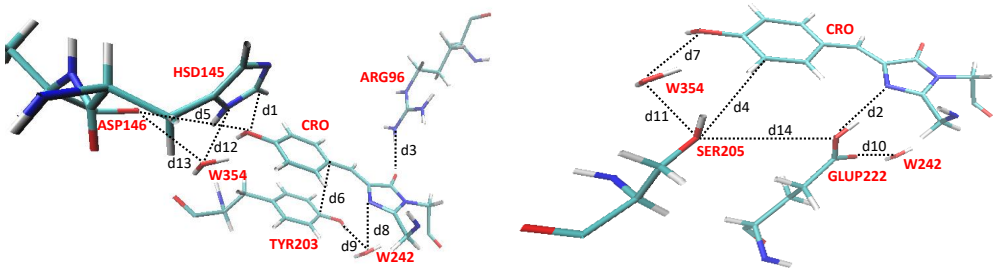


FIG. 6: Definitions of selected distances used to compare various structures for the ON-form: d1 = CRO:OH-HIS145:CE1, d2 = CRO:N2-GLU222:OE1, d3 = CRO:O2-ARG96:NH2, d4 = CRO:CE2-SER205:OG, d5 = CRO:OH-ASP146:O, d6 = CRO66:CG2-TYR203:CZ, d7 = CRO66:OH-TIP354:OH2, d8= CRO:N2-TIP242:O,d9= TYR203:OH-TIP242:OH2, d10= GLU222:OE2-TIP242:OH2, d11 = SER205:OG-TIP354:OH2, d12 = HIS145:ND1-TIP354:OH2, d13 = ASP146:O-TIP354:OH2, d14= GLU222:OE1-SER205:OG.

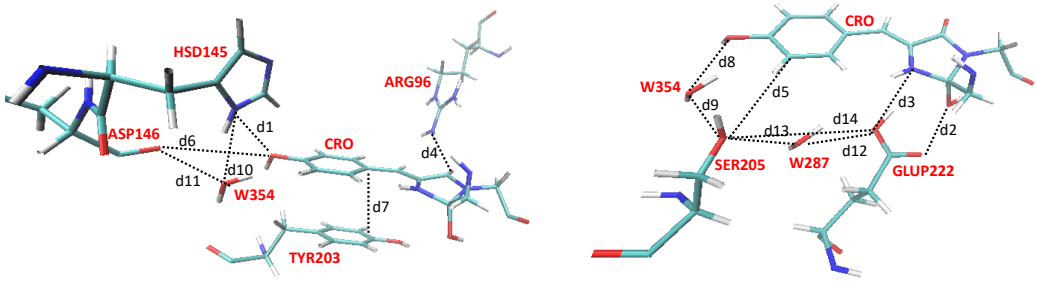


FIG. 7: Definitions of selected distances used to compare various structures for the OFF-form: d1 = CRO:OH-HIS145:CE1, d2 = CRO:O1-GLU222:OE1, d3 = CRO:N2-GLU222:OE2, d4 = CRO:O2-ARG96:NH2, d5 = CRO:CE2-SER205:OG, d6 = CRO:OH-ASP146:O, d7 = CRO66:CG2-TYR203:CZ, d8 = CRO66:OH-TIP245:OH2, d9 = SER205:OG-TIP245:OH2, d10 = HIS145:ND1-TIP245:OH2, d11 = ASP146:O-TIP245:OH2, d12= GLU222:OE2-TIP287:OH2, d13= SER205:OG-TIP287:OH2, d14= GLU222:OE1-SER205:OG.

ment (as judged from the smallest standard deviations of the QM/MM optimized structures from the X-ray structure) for the HSE-GLUP state (this is in agreement with Ref. 20). The largest variations between different protonation states are observed for d2 (Glu222-imidozalinone) and d14 (Ser205-Glu222). For the latter, the crystal structure value is 4.18 Å and the HSE-GLUP value is 4.72 Å, whereas other protonation states yield shorter distances — e.g., in the structures with GLU d14∼2.5 Å. In Ref. 20, d18=4.87 Å for HSE-GLUP

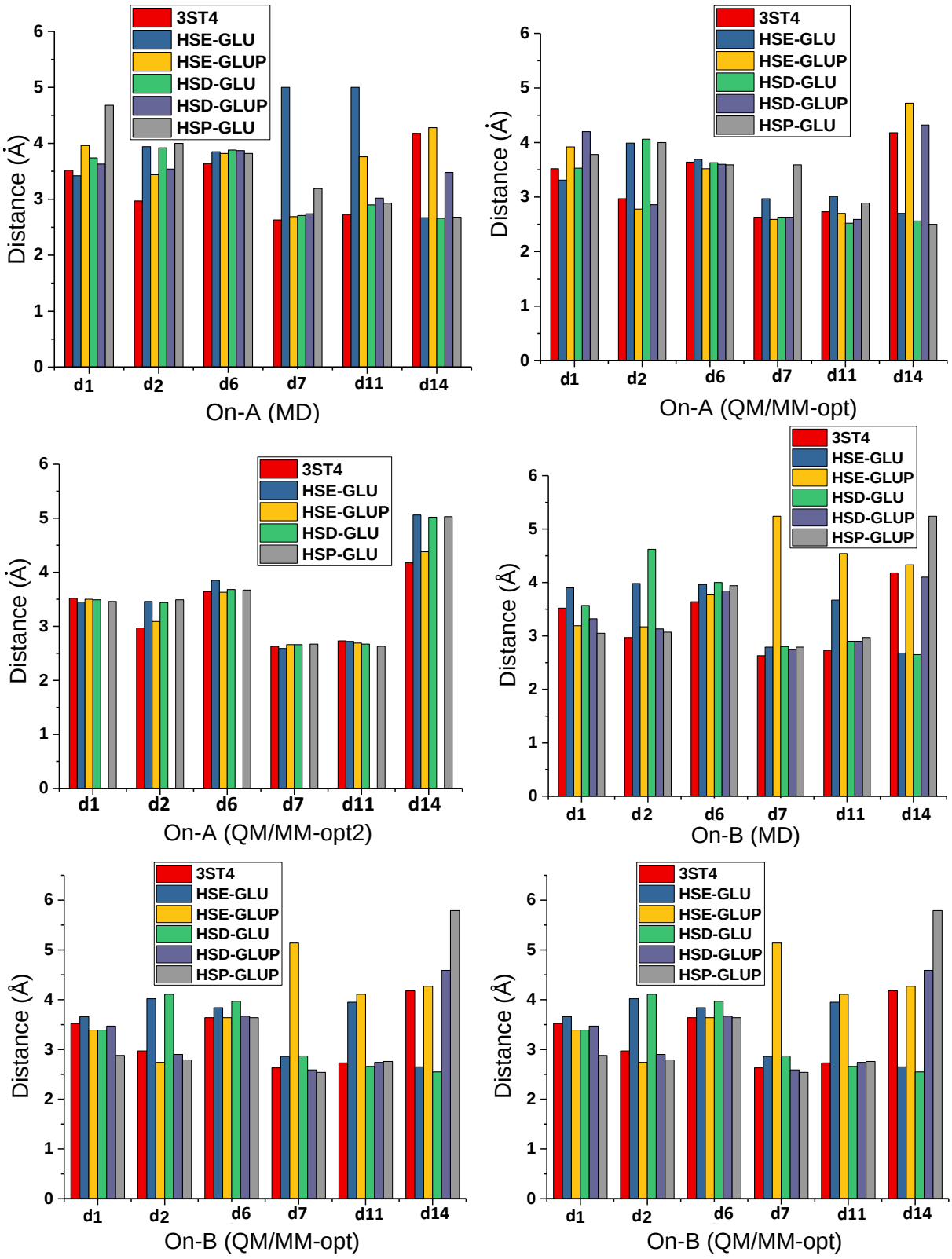


FIG. 8: Key distances for ON-states: Comparison between crystal structure, average MD values, and QM/MM optimization. See Fig. 6 for definitions.

(neutral ON-state), which is close to our present value. Further comparison between the present model structures and those from Ref. 20 are available in the SI (Tables S7 and S8 and Figure S4).

In terms of the total electronic energies, HSE-GLUP is 0.33 eV below HSD-GLUP, which is 0.42 eV lower than HSP-GLU; this energetics is consistent with the HSE-GLUP state being the most favorable for the ON-state with the neutral chromophore. The gap between HSD-GLU and HSE-GLUP is 1.05 eV.

For the anionic chromophore, we observe the best agreement for HSD-GLUP. Here again d2 and d14 show the largest variations between different protonation states. The HSE-GLUP state is also a viable candidate. In contrast, in Ref. 20 HSP-GLUP was used to describe anionic ON-state. In terms of the structures, the largest difference between HSD-GLUP and HSP-GLUP is in d14: compare 4.18 Å (X-ray) with 4.59 Å (HSD-GLUP) and 5.79 Å (HSP-GLUP). For HSE-GLUP, the largest differences are observed for d11 (Wat-Asp146) and for Chro-Tyr203: compare 5.14 Å (in HSE-GLUP) versus 2.59 (HSD-GLUP). Here again, MD simulations and QM/MM optimizations are in qualitative agreement.

In terms of the total electronic energies, HSD-GLUP is only 0.35 eV below HSE-GLUP. HSD-GLU and HSE-GLU are nearly isoenergetic (the latter is 0.1 eV lower). Hence, on the basis of structures and energetics, HSD-GLUP appears to be the best match, but other states cannot be ruled out.

Fig. 9 shows the key distances for the OFF-state. Here the differences between the MD values and QM/MM optimizations are much larger (some values are off the chart), highlighting the advantage of using rigorous QM potentials. For the OFF-state, we observe the best agreement in terms of structures for the HSE-GLUP2, HSD-GLU, and HSP-GLU, but the differences are not that large. Comparisons of the total energies favor HSD-GLUP-OE2 (among HSD-GLUP-OE2, HSP-GLU, HSE-GLUP-OE2, HSE-GLUP, and HSD-GLUP series) and HSD-GLU (relative to HSE-GLU); HSP-GLU is slightly most stable compared to HSE-GLUP-OE2 (0.09 eV). In Ref. 20, HSE-GLUP and HSP-GLU were chosen as the best candidates.

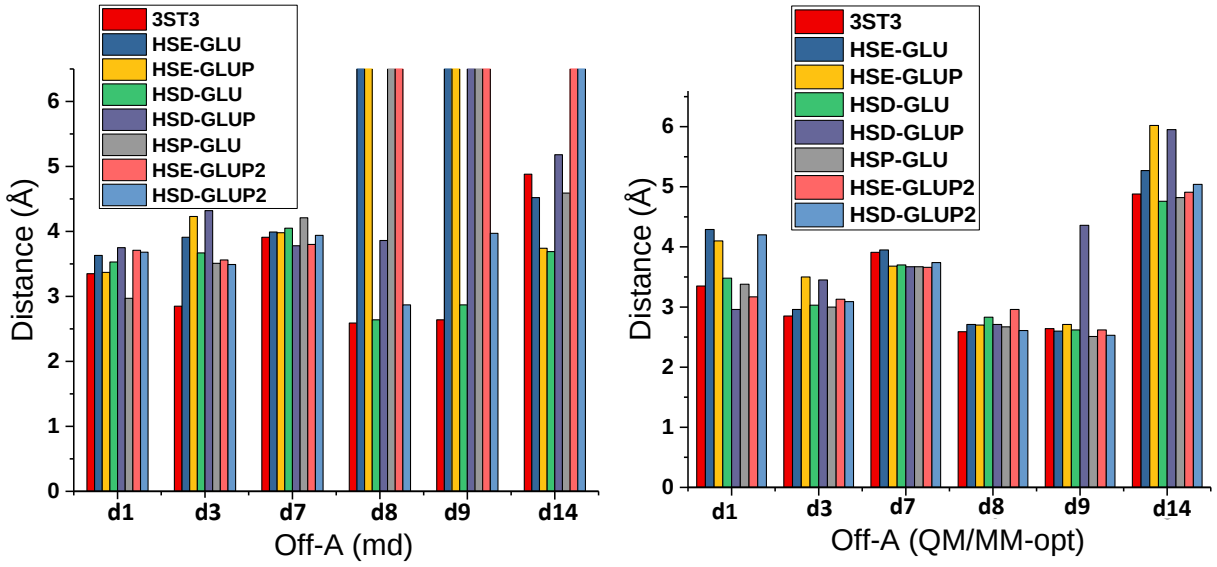


FIG. 9: Key distances for OFF-states: Comparison between crystal structure, average MD values, and QM/MM optimizations. See Fig. 7 for definitions. Note that some MD values for d8, d9, and d14 are off the chart.

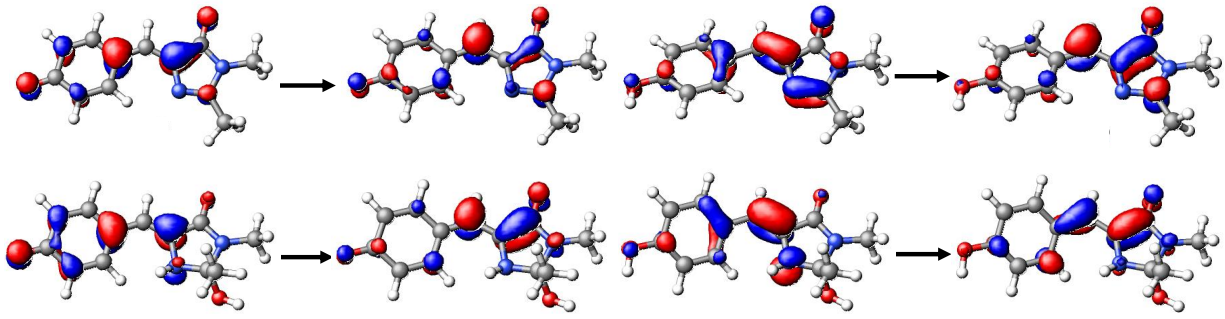

FIG. 10: NTOs for the lowest bright states of the bare chromophores. Top left: anionic ON-state; top right: neutral ON-state; bottom left: anionic OFF-state; bottom right: neutral OFF-state. $\omega$B97X-D/aug-cc-pVDZ.

### B. Excited-state analysis

We begin by analyzing excited states of the isolated chromophores computed at their equilibrium geometries (see the SI). Table I shows computed excitation energies and oscillator strengths of the isolated chromophores in ON- and OFF-states and Fig. 10 shows the respective natural transition orbitals (NTOs). The excited state of a GFP-like chromophore

TABLE I: Excitation energies (eV) of the isolated chromophores (ON- and OFF-states, A and B forms) computed at the optimized geometries ($\omega$B97X-D/aug-cc-pVDZ). Oscillator strengths are shown in parenthesis[a].

| System | TDDFT aug-cc-pVDZ | SOS-CIS(D) aug-cc-pVDZ | EOM-CCSD aug-cc-pVDZ | XMCQDPT2 cc-pVDZ | XMCQDPT2 aug-cc-pVDZ-mod[b] |
|---|---|---|---|---|---|
| on-A | 3.75 (0.72) 331 nm | 3.88 (1.04) 320 nm | 3.69 (0.98) 336 nm | 3.54 (0.50) 350 nm | 3.26 (0.49) 380 nm |
| on-B | 3.10 (1.00) 400 nm | 2.75 (1.07) 451 nm | 2.95 (1.14) 420 nm | 2.58 (1.11) 481 nm | 2.40 (1.02) 517 |
| off-A | 4.29 (0.60) 289 nm | 4.62 (0.71) 268 nm | 4.01 (0.88) 309 nm | 4.47 (0.68)[c] 277 nm | 4.04 (0.42) 307 nm |
| off-B | 3.39 (0.91) 366 nm | 3.07 (1.07)[d] 404 nm | 3.23 (1.11) 384 nm | 3.02 (0.90) 411 nm | 2.82 (0.78) 440 nm |

[a] The lowest excited state is the bright state in all cases except when marked otherwise.

[b] aug-cc-pVDZ on heavy atoms and cc-pVDZ on hydrogens.

[c] The lowest bright state corresponds to the $S_0$-$S_2$ transition.

[d] The lowest bright state corresponds to the $S_0$-$S_2$ transition.

corresponds to the $\pi \rightarrow \pi^*$ transition, with the main action happening on the methyne bridge[43]. Consistently with previous studies[44,45], we observe that lower-level methods (TD-DFT) overestimate the excitation energies. EOM-CCSD energies are 0.06-0.28 eV below the TD-DFT ones. XMCQDPT2 energies are below the EOM-CCSD energies by 0.4-0.5 eV for on-A and on-B, but are nearly he same for off-A. We note a generally good agreement between SOS-CIS(D) and XMCQDPT2 for all four cases: the differences are less than 0.4 eV and XMCQDPT2 values are below SOS-CIS(D). However, all methods capture (qualitatively) the large red shift ($\sim$0.6 eV) between the neutral and anionic chromophores (we note that SOS-CIS(D) overestimates the shift by almost a factor of 2). The shift can be explained in the framework of the Hückel model[43] and is due to the increased delocalization on the methine bridge in the anionic form. The oscillator strength for the anionic form is higher than that for the neutral, but the values depend on the method, i.e., the ratio is 1.2 for EOM-CCSD, 1.4 for TD-DFT, and 2.2 for XMCQDPT2.

As expected, the excitation energies in the hydrated chromophore (OFF-state) are blue-shifted relative to the ON-state by roughly 0.6 eV due to disrupted conjugation. Here again all methods are in a qualitative agreement, although SOS-CIS(D) yields much higher values than TD-DFT, EOM-CCSD, and XMCQDPT2.

As the next step, we consider the effect of the environment on excitation energies. The protein environment is important for quantitative comparison of the theoretical values with the experiment[12,34–36]. Here we primarily rely on a finite cluster approach and compute excitation energies using the extended QM system (Fig. 5). To assess the effect of the protein beyond the extended QM, we also include the results of the QM/MM calculations using electrostatic embedding computed for 21 snapshots taken from the MD simulations.

Figures S7 and S8 in the SI show the NTOs (TD-DFT/aug-cc-pVDZ/extended QM) of the lowest states of the neutral and anionic forms for the ON state. As one can see, the two lowest states in the on-A form correspond to the CT state (Tyr203→Chro) and bright LE state ($\pi\pi^*$). In the on-B form, the lowest excited state of $\pi\pi^*$ character. In QM-only calculations (extended QM), the second excited state is of CT character (0.3-0.8 eV above), but this states disappear when the rest of the protein is included. The protonation states of His145 and Glu222 affect the excitation energies, but not the characters of the states. Importantly, the low-lying CT state appear in all protonation forms of on-A.

Table S14 in the SI shows TD-DFT excitation energies computed for large and extended QM with different basis sets for the on-A form. As one can see, protein environment (compare extnednded QM with bare chromophore) leads to a red shift of the excitation energy of the LE state by 0.2-0.4 eV. We observe the lowest excitation energy HSE-GLUP (the most likely protonation state) and the highest in HSP-GLU. The differences between large and extended QM are less than 0.1 eV. The effect of the basis set is small — for all forms, changing the basis from the cc-pVDZ to a mixed basis (aug-cc-pVDZ on the chromophore and Tyr203 and cc-pVDZ on the rest) and to the full aug-cc-pVDZ basis leads to small red shifts for all protonation states; the largest magnitude was 0.06 eV. To estimate the effect of the rest of the protein, we compare the results of the QM and QM/MM calculations using MD snapshots (Table S13 in the SI): as one can see, including the rest of the protein leads to a small blue shift of about 0.1 eV for the LE state. The results for the CT state show somewhat stronger dependence on the computational protocol. At the TD-DFT level, the CT state appears 0.3-0.5 eV below the LE state in finite-cluster calculations. Increasing

basis set can blue-shift its energy by up to 0.03 eV. Interestingly, including the effect of the rest of protein (Table S13) leads to a larger blue-shift of the CT state than for the LE state (0.25 versus ∼0.1). The results suggest that the position of the CT state in the QM-only calculations is slightly underestimated. We attribute this effect to the overstabilization of the CT state by the positively charged arginine in finite-cluster calculations; including the rest of the protein and the counterions leads to the partial screening of the arginine field and, therefore, increases the energy of the CT state. We also note that the position of the CT state is sensitive to the counterions and varies among the snapshots; this is similar to the trends reported in Refs. 34,46.

Importantly, even including this additional correction, the CT state appears below the LE state at TD-DFT level in the neutral chromophore in all protonation states of His145 and Glu222. To further refine the positions of the LE and CT states, we computed excitation energies using XMCQDPT2; these results are collected in Table S17. Similar to the isolated chromophore, the XMCQDPT2 excitation energies of the $\pi\pi^*$ state are red-shifted relative to TD-DFT. The inclusion of the protein environment has the same effect as in TD-DFT — overall red shift relative to the isolated chromophore. In the finite-cluster calculations, the XMCQDPT2 excitation energies of the LE state appear to be red-shifted relative to the experiment by 0.4 eV in HSD-GLUP and by 0.2 in HSE-GLUP protonation states; including the effect of the rest of the protein is expected to reduce this discrepancy. Importantly, XMCQDPT2 calculations confirm the presence of the CT state. At this level of theory, the gap between the LE and CT states is smaller than at TD-DFT level, which is consistent with the tendency of TD-DFT to overestimate the positions of valence excited states and to underestimate the position of CT states. In the HSE-GLUP form, our best candidate for the neutral ON-form, the CT state is 0.3 eV below LE state at XMCQDPT2 state in finite-cluster calculation. Extrapolating to the full protein, we expect this gap to shrink to about 0.15 eV.

These comparisons provide a measure of the uncertainty of the calculations of the due to basis set, QM size, and the correlation treatment; they also quantify the variations due to different protonation states. Importantly, although we cannot pin the exact location of the CT state, our results indicate that it is energetically close to the LE state. Taking into account the variations in energies due to different protonation states and uncertainties of computational protocols, we estimate that the CT state is within 0.25 eV from the LE state

in the neutral ON-state. We also observe that its position is very sensitive to the hydrogen bond pattern and the dynamics of counter-ions. Hence, its energy can fluctuate in the course of thermal motions, bringing it in resonance or below the LE state. Hence, the CT state can be accessed either via direct excitation (since it has non-zero oscillator strength) or via non-adiabatic transition from the LE state. Since CT states are known to be involved in bleaching and some photochonversions[36,46–48], the appearance of this state in Dreiklang is highly suggestive of its role in photoconversion. Below we further investigate this question.

Tables S15 and S18 show the results for the anionic chromophore (on-B form). In this case, all methods (TD-DFT and XMCQDPT2, both finite-cluster and QM/MM calculations) agree that the lowest state is LE of $\pi\pi^*$ character. In finite-cluster calculations, TD-DFT shows the CT state about 0.3-0.6 eV above the LE state, but when the rest of the protein is included, this state disappears. The effect of the protein leads to a small shift of the LE state (-0.2/+0.02 eV). The differences between the cc-pVDZ and aug-cc-pVDZ bases do not exceed 0.1 eV. Better treatment of electron correlation leads to substantial red-shift, up to 0.6 eV. Comparing to the experimental value (2.43 eV), the XMCQDPT2 values are within 0.1-0.2 eV, depending on the protonation state. At the XMCQDPT2 level, the best agreement is observed for HSE-GLUP and HSD-GLU structures.

The results for the OFF-state (shown in Tables S16 and S19 in the SI) reveal similar trends. In this case, the protein environment leads to larger red shifts of 0.4-0.8 eV, depending on the protonation state. On the XMCQDPT2 level, the best agreement with experiment is observed for HSE-GLUP and HSE-GLUP2 structures (and the largest deviation — for HSE-GLU and HSD-GLUP2).

In summary, Fig. 11 shows these results for excited-state calculations graphically. Based on our results, the excitation energies in different protonation states are close and cannot be used to rule out some protonation states. Moreover, given the small energy differences between the respective optimized structures, different protonation states can be populated simultaneously. The excitation energies for our best candidates (based on the structural analysis) are as follows: on-A/HSE-GLUP — TD-DFT is 3.4 eV and XMCQDPT2 is 2.93 eV (including extrapolation), to be compared with 3.01 experimental value; on-B/HSD-GLUP — TD-DFT is 3.0 eV and XMCQDPT2 is 2.2 eV, to be compared with 2.43 eV experimental value; on-B/HSE-GLUP — TD-DFT is 3.0 eV and XMCQDPT2 is 2.4 eV, to be compared with 2.43 eV experimental value; and for the off-A/HSE-GLUP2 form — 3.9

eV/3.6 eV; HSD-GLU: 3.7/3.4 eV, HSP-GLU: 3.9/3.5 eV; all these numbers are reasonably close to the experimental value of 3.65 eV. The results also suggest that the shoulder at 2.58 eV in the absorption spectrum of the ON-state (Fig. 2) may be due to either the presence of another major protonation state or CT state of the neutral chromophore.

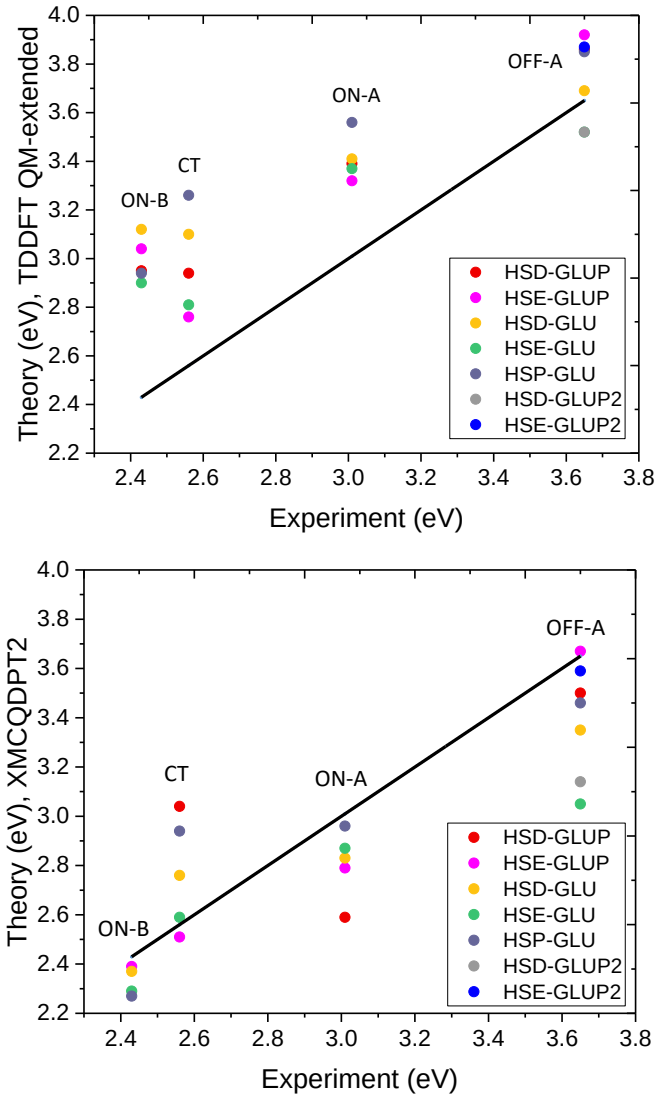


FIG. 11: Excitation energies for different model system shown against the experimental values. Top: TD-DFT/aug-cc-pVDZ; bottom: XMCQDPT2/aug-cc-pVDZ/cc-pVDZ. Extended QM.

### C. Implications of the CT state and possible mechanism for photo-reaction

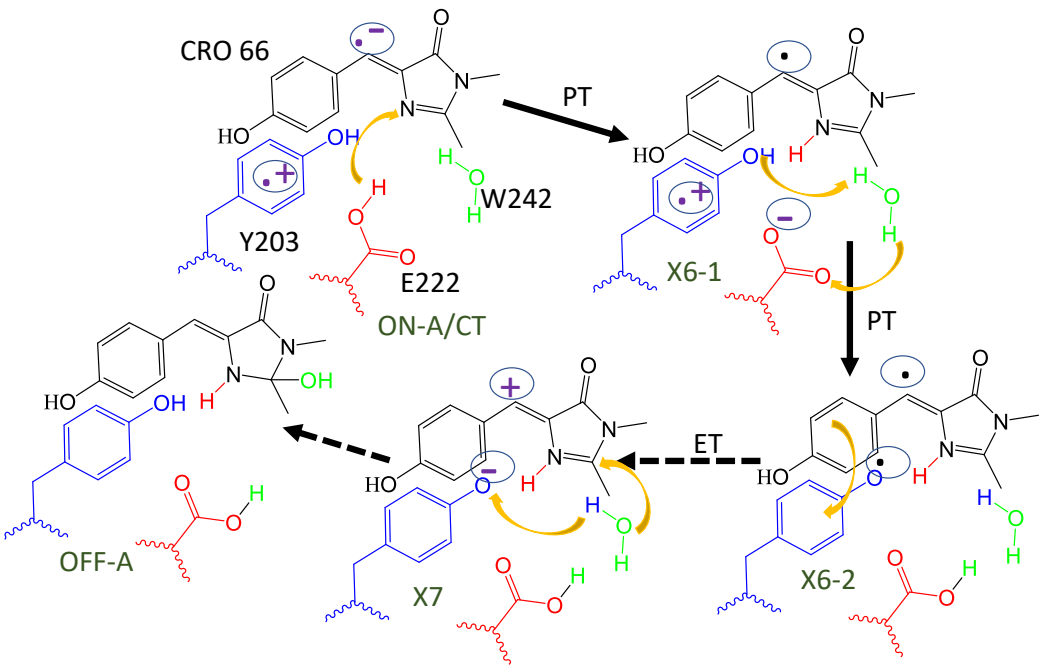


FIG. 12: Proposed reaction on the CT state. AIMD and excited-state optimization reveal that the steps leading to the formation of X6-2/X7 are nearly barrierless and proceed on the scale of ∼100-200 fs. The last two steps (shown by dashed arrows), back electron transfer from Chro to Tyr203, nucleophilic addition of $OH^-$ to Chro, and reprotonation of Tyr203, are hypothesized. The structures of the possible intermediates are defined in Fig. S9 in the SI.

Fig. 4 shows the essential steps of Dreiklang photocycle and outlines the revised photoconversion mechanism via the CT state. The CT state can be populated either via direct excitation or by non-adiabatic transition from the LE state. This is followed by a rapid proton transfer from a nearby residue. The protonated neutral radical chromophore looses the extra electron and undergoes nucleophilic addition of $OH^-$ from the nearby water. Below we describe computational support of the proposed mechanism.

To investigate possible excited-state pathways, we carried out geometry optimization and AIMD simulations for the CT and LE states (for the on-A-HSE-GLUP structure). Fig. 12 shows the structural transformation along the AIMD/optimization trajectories. Fig. 13 shows additional details of the AIMD simulations: energy profiles of the two lowest electronic states (Kohn-Sham reference state and the lowest TD-DFT state) and the charges of the

key residues (chromophore and Tyr203) in these two states; the results are shown in the right panel of Fig. 13. The abrupt changes in the charges clearly indicate the instances of proton transfer.

In the CT state, both the optimization and AIMD simulations show rapid (on the scale of ∼100-250 fs) and barrierless proton-transfer steps leading to the formation of the protonated chromophore; this can be rationalized by increased basisity of the imidozalinone nitrogen caused by the electron attachment. First, proton is transferred from Glu222 to imidozalinone nitrogen (this happens within 50 fs). Then Glu222 is reprotonated via proton transfer from Tyr203 (acidified as a result of the electron transfer to the chromophore) via water-mediated pathway. This process is completed in 200-250 fs. At this point, the CT state is energetically nearly degenerate with the reference Kohn-Sham state ($S_0$); or, in other words, the Chro-Tyr203 radical pair (neutral protonated radical chromophore and neutral deprotonated Tyr203 radical, X6-2 structure) is nearly isoenergetic with the closed-shell ion-pair state (in which the chromophore is protonated and positively charged and Tyr203 is deprotonated and negatively charged, X7 structure). Hence, one can assume back-electron transfer resulting in the formation of the ground-state X7 intermediate.

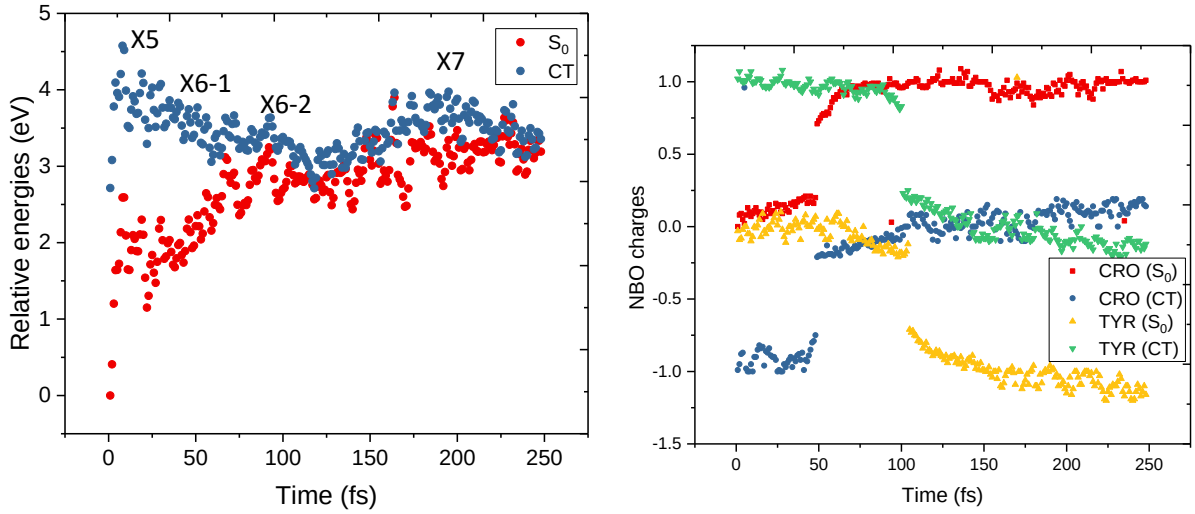


FIG. 13: Left: Energies of the Kohn-Sham reference state ($S_0$) and CT state along the AIMD trajectory on the CT potential energy surface. Right: Charges on the chromophore and Tyr203 in the Kohn-Sham reference state and the CT state (lowest TDDFT state). Labels X5, X6, and X7 denote points along the trajectories when structures resembling these intermediates are formed (see Fig. S9 in the SI; X6-1 refers to HSE-GLU; X6-2 refers to HSE-GLUP2).

Based on these observations, X6-2 (Chro$^{\cdot}$-Tyr203$^{\cdot}$ radical pair) or X7 (Chro$^{+}$-Tyr203$^{-}$ zwitterion) are our candidates for the intermediate X observed spectroscopically in time-resolved study[10]. Experimentally[10], strong transient absorption was observed at 2.67-2.88 eV (with 100 fs kinetics); then at the nanosecond scale formation of the intermediate X adsorbing at 2.76 eV was observed. Hence, both short-time transient absorption and longer time-scale absorption occurs at about 2.8 eV, which is 0.2 eV red-shifted relative to the absorption of the form A.

At the geometry taken from AIMD trajectory at time ∼248 fs, the excitation energy of the zwitter ion X7 (computed as the lowest bright transition from the Kohn-Sham reference state) is 3.56 eV (oscillator strength 0.40), which is too high compared to the experimental absorption of X. On the other hand, the excitation energy of the radical pair X6-2 (computed as the lowest bright transition from the lowest TDDFT state) is 2.93 eV (oscillator strength 0.22). The large difference in the excitation energies of the two structures can be easily rationalized: protonation of the closed-shell neutral chromophore should lead to a blue-shift relative to the parent on-A form, whereas the absorption of the radical anion (chromophore with the additional electron) or protonated neutral radical is expected to be red-shifted relative to the respective closed-shell parent species. Hence, X6-2 appears to be a good candidate for the hot I$^{*}$ intermediate formed on the femtosecond timescale[10]. Further changes in its excitation energy (leading to small blue shift in the absorption of X relative to I$^{*}$) might be due to the structural relaxation of the protein.

Back electron-transfer step (Chro→Tyr203) would result in the formation of X7 in which the chromophore is positively charged and, therefore, appears to be a good candidate for nucleophilic attack by the nearby water, leading to the formation of hydrated chromophore and reprotonation of Tyr203. Our preliminary calculations indicate that this step would require to overcome a barrier — the scan along water-imidozalinone distance (Fig. S14 in the SI) yields a barrier of ∼25 kcal/mol, which is very similar to the barrier of the thermal recovery reaction. This is relatively crude estimate, which should be regarded as an upper bound on the barrier; more accurate estimates would be a subject of future studies. The delayed appearance of the hydrated chromophore is consistent with such a barrier. We note that, in contrast to thermal recovery reaction, the reaction may still be rather fast, because of the high excess energy available to the system (see left panel in Fig. 13). We validated (by AIMD simulations) that once the system reaches this transition state, the dynamics swiftly

proceeds downhill, leading to the formation of the hydrated chromophore and reprotonated Tyr203. The AIMD simulations also show that the reverse reaction, from X7 back to the neutral chromophore and reprotonated Tyr203, is very efficient and can compete with the final step of the nucleophilic addition.

In contrast to the CT state, geometry optimization and AIMD simulation (1 ps long trajectory) on the LE potential energy surface do not show any significant structural changes, i.e., no evidence of the ultrafast ESPT from the chromophore (Fig. S13) posited in Ref. 10.

In summary, the following picture of Dreiklang's photocycle emerges from the results of our theoretical modeling:

1. Excitation of the anionic form (peak B in the ON-state) leads to fluorescence.

2. Excitation of the neutral form (peak A in the ON-state) leads non-adiabatic transition to the CT state, from which photochemical transformation is ensued. It can also lead to ESPT and fluorescence from the anionic state (as in main photocycle of wt-GFP), but this channel is strongly suppressed.

This picture differs from the mechanism outlined in Ref. 10, where it was proposed that photochemistry unravels in the anionic state, formed by ESPT from photoexcited form A. An obvious flaw in the ESPT mechanism is that it does not explain why there is no photoconversion upon the direct excitation of the anionic form. In contrast, our proposed mechanism via the CT state, which can only be populated by the excitation of the neutral form, explains the essential trait of Dreiklang, the decoupling of the fluorescence excitation (produced via the anionic form) from the photoconversion (produced by excitation of the neutral form).

Ref. 10 invoked ESPT because of the observed isotope effect. This effect can be explained by concerted proton transfer to the chromophore in CT state. Ref. 10 postulated ESPT to explain the observed short-time dynamics (510 fs) and commented out that this process is an order of magnitude faster than ESPT in GFP (2 ps), which is inconsistent with lack of strong fluorescence following the excitation of peak A and increased pKa of the chromophore. Our AIMD simulations on the LE PES show no evidence of the ultrafast ESPT. The authors of Ref. 10 also commented that the ESPT dynamics in Dreiklang is significantly less sensitive to H/D exchange than in GFP, deuterium slowing the observed kinetics by a factor of 1.5 instead of 5. Our simulations strongly suggest that what is seen on the femtosecond time

scale is formation of radical pair Chro$^{\cdot}$-Tyr203$^{\cdot}$ in which the chromophore is protonated on imidozalinone's nitrogen and Tyr203 is deprotonated. Our dynamics shows 250 fs time for proton transfers, but one needs also to include time for non-adiabatic transition from the LE state populating the CT state. Sub-picosecond timescales are very likely and there should be some isotope effect.

In addition, our revised photocycle is consistent with the following observations. As pointed out in Ref. 7, the essential difference between the parent system (Citrine) and Dreiklang is the upshift of the pKa of the ON-state of the chromophore (7.2 versus 5.7), which increased the effectiveness of photoconversion. Larger pKa suggests that the ESPT from the neutral form is suppressed, making the population of CT state more competitive. Note that fluorescence excitation spectrum (Fig. 1B from Ref. 7) shows that very little fluorescence is produced by excitation of the peak A. We note that the photoconversion is achieved by continuous irradiation in the course of $\sim$5 s, which suggests that the quantum yield for this process is relatively small.

Ref. 7 emphasized that Tyr203 and Glu222 (and Gly65) are crucial for Dreiklang function. They also comment that in the fluorescent-state, Tyr203 and Glu222 form hydrogen bonds to a water molecule and thereby stabilize it in close vicinity to the C65 of the chromophore, a situation that is different in the nonswitchable GFP (avGFP-S65T). This strengthens the argument that the reaction may proceed by concerted proton transfer from water-to-Glu222-to-Chro.

We conclude by noting that neither the ET step (population of the CT state) nor the subsequent the barrierless proton-transfer steps should be affected by temperature, meaning that these steps would not be suppressed at cryogenic temperature. We also note that previous studies indicate that the photoinduced recovery of the ON-state might be barierrless[20]. This suggest that Dreiklang could be a good starting point for developing photoswitchable fluorescent protein that can operate at low temperatures, as desired for cryogenic super-resolution imaging applications[49,50].

## IV. CONCLUSION

In this contribution, we investigated properties of the fluorescent protein Dreiklang using high-level electronic structure methods combined with QM/MM and dynamics simulations.

The results allowed us to quantify the spectral consequences of possible protonation states of the key residues around the chromophore and to refine the properties of the low-lying excited states. The key finding is that the neutral (protonated) ON-state of Dreiklang features low-lying state of CT character (Tyr203→Chro), which is energetically close to the LE and is stongly affected by hydrogen bonding and thermal motions. Once this state is populated (either by direct photoexcitation or via non-adiabatic transition), the system undergoes a cascade of proton transfer steps leading to the protonation of the chromophore (on imidozalinone's nitrogen) and formation of the neutral Chro-Tyr203 radical pair, nearly iso-energetic with the zwitter-ionic state (in which Tyr203 is in deprotonated anionic state and the chromophore is positively charged). This structure appears to be a good candidate for nucleophilic addiction of hydroxide to the chromophore, coupled with reprotonation of Tyr203.

This mechanism is consistent with the available experimental data. The disrupted hydrogen-bonding network around the chromophore and its reduced acidity explain that the canonical ESPT route is strongly suppressed, making the CT channel competitive. The key role of the CT state, which is only accessible by photoexcitation of the on-A form, explains the unique feature of Dreiklang, the decoupling of fluorescence from the photoswitching.

## Acknowledgments

This work was supported by the U.S. National Science Foundation (No. CHE-1856342). YM was supported by China Scholarship Council during her stay at USC. IVP and BLG thank the Russian Science Foundation (17-13-01051) for the financial support of this study. IVP, BLG, and AVN acknowledge the use of computing facilities of supercomputer resources of the Joint Supercomputer Center of the Russian Academy of Sciences and the equipment of the shared research facilities of HPC computing resources at the Lomonosov Moscow State University.

YM present address is Inner Mongolia University of Technology, Hohhot, P.R. China.

**Supporting information is available:** Definitions of the key structures, computational details, forcefield parameters for the OFF-state, detailed structural analysis, excitation en-

ergies, NTOs.

## Conflict of interest

A.I.K. is the president and a part-owner of Q-Chem, Inc.

---


[1] A. Acharya, A. M. Bogdanov, K. B. Bravaya, B. L. Grigorenko, A. V. Nemukhin, K. A. Lukyanov, and A. I. Krylov, Photoinduced chemistry in fluorescent proteins: Curse or blessing?, Chem. Rev. **117**, 758 (2017).

[2] K. Nienhaus and G.U. Nienhaus, Fluorescent proteins for live-cell imaging with super-resolution, Chem. Soc. Rev. **43**, 1088 (2014).

[3] H.C. Ishikawa-Ankerhold, R. Ankerhold, and G.P.C. Drummen, Advanced fluorescent microscopy techniques — FRAP, FLIP, FLAP, FRET and FLIM, Molecules **17**, 4047 (2012).

[4] R. Ando, H. Mizuno, and A. Miyawaki, Regulated fast nucleocytoplasmic shuttling observed by reversible protein highlighting, Science **306**, 1370 (2004).

[5] M. Andresen, A.C. Stiel, J. Fölling, D. Wenzel, A. Schönle, A. Egner, C. Eggeling, S.W. Hell, and S. Jakobs, Photoswitchable fluorescent proteins enable monochromatic multilabel imaging and dual color fluorescence nanoscopy, Nat. Biotechnol. **26**, 1035 (2008).

[6] D.M. Chudakov, A.V. Feofanov, N.N. Mudrik, S. Lukyanov, and K.A. Lukyanov, Chromophore environment provides clue to "kindling fluorescent protein" riddle, J. Biol. Chem. **278**, 7215 (2003).

[7] T. Brakemann, A.C. Stiel, G. Weber, M. Andresen, I. Testa, T. Grotjohann, M. Leutenegger, U. Plessmann, H. Urlaub, C. Eggeling, M.C. Wahl, S.W. Hell, and S. Jakobs, A reversibly photoswitchable GFP-like protein with fluorescence excitation decoupled from switching, Nat. Biotechnol. **29**, 942 (2011).

[8] R. M. Wachter, M. A. Elsliger, K. Kallio, G. T. Hanson, and S. J. Remington, Structural basis of spectral shifts in the yellow-emission variants of green fluorescent protein, Structure **6**, 1267 (1999).

[9] M. Ormö, A.B. Cubitt, K. Kallio, L.A. Gross, R.Y. Tsien, and S.J. Remington, Crystal structure of the aequorea victoria green fluorescent protein, Science **273**, 1392 (1996).

[10] F. Lacombat, P. Plaza, M.-A. Plamont, and A. Espagne, Photoinduced chromophore hydration in the fluorescent protein Dreiklang is triggered by ultrafast excited-state proton transfer coupled to a low-frequency vibration, J. Phys. Chem. Lett. **8**, 1489 (2017).

[11] R.Y. Tsien, The green fluorescent protein, Annu. Rev. Biochem. **67**, 509 (1998).

[12] B. L. Grigorenko, A. V. Nemukhin, I. V. Polyakov, D. I. Morozov, and A. I. Krylov, First-principle characterization of the energy landscape and optical spectra of the green fluorescent protein along A-I-B proton transfer route, J. Am. Chem. Soc. **135**, 11541 (2013).

[13] S.R. Meech, Excited state reactions in fluorescent proteins, Chem. Soc. Rev. **38**, 2922 (2009).

[14] J.J. van Thor, Photoreactions and dynamics of the green fluorecent protein, Chem. Soc. Rev. **38**, 2935 (2009).

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

[16] K. Brejc, T.K. Sixma, P.A. Kitts, S.R. Kain, R.Y. Tsien, M. Ormö, and S.J. Remington, Structural basis for dual excitation and photoisomerization of the Aequorea victoria green fluorescent protein, Proc. Nat. Acad. Sci. **94**, 2306 (1997).

[17] S.J. Remington, Green fluorescent protein: A perspective, Protein Sci. **20**, 1509 (2011).

[18] R. Simkovitch, A. Huppert, D. Huppert, S.J. Remington, and Y. Miller, Proton transfer in wild-type GFP and S205V mutant is reduced by conformational changes of residues in the proton wire, J. Phys. Chem. B **117**, 11921 (2013).

[19] B. L. Grigorenko, A. I. Krylov, and A. V. Nemukhin, Molecular modeling clarifies the mechanism of chromophore maturation in the green fluorescent protein, J. Am. Chem. Soc. **139**, 10239 (2017).

[20] B. L. Grigorenko, I. Polyakov, A. I. Krylov, and A. V. Nemukhin, Computational modeling reveals the mechanism of fluorescent state recovery in the reversibly photoswitchable protein Dreiklang, J. Phys. Chem. B **123**, 8901 (2019).

[21] M. H. M. Olsson, C. R. Sondergaard, M. Rostkowski, and J. H. Jensen, PROPKA3: Consistent treatment of internal and surface residues in empirical pKa predictions, J. Chem. Theory Comput. **7**, 525 (2011).

[22] N. Foloppe and A. D. MacKerell, All-atom empirical force field for nucleic acids: I. Parameter optimization based on small molecule and condensed phase macromolecular target data, J.

Comput. Chem. **21**, 86 (2000).

[23] N. Reuter, H. Lin, and W. Thiel, Green fluorescent proteins: Empirical force field for the neutral and deprotonated forms of the chromophore. Molecular dynamics simulations of the wild type and S65T mutant, J. Phys. Chem. B **106**, 6310 (2002).

[24] J.-D. Chai and M. Head-Gordon, Systematic optimization of long-range corrected hybrid density functionals, J. Chem. Phys. **128**, 084106 (2008).

[25] J.-D. Chai and M. Head-Gordon, Long-range corrected hybrid density functionals with damped atom-atom dispersion interactions, Phys. Chem. Chem. Phys. **10**, 6615 (2008).

[26] Y. Zhang and W. Yang, A challenge for density functionals: Self-interaction error increases for systems with a noninteger number of electrons, J. Chem. Phys. **109**, 2604 (1998).

[27] V. Polo, E. Kraka, and D. Cremer, Electron correlation and the self-interaction error of density functional theory, Mol. Phys. **100**, 1771 (2002).

[28] M. Lundber and P.E.M. Siegbahn, Quantifying the effects of the self-interaction error in DFT: When do the delocalized states appear?, J. Chem. Phys. **122**, 224103 (2005).

[29] A. Dreuw, J.L. Weisman, and M. Head-Gordon, Long-range charge-transfer excited states in time-dependent density functional theory require non-local exchange, J. Chem. Phys. **119**, 2943 (2003).

[30] S. Grimme, Accurate description of van der Waals complexes by density functional theory including empirical corrections, J. Comput. Chem. **25**, 1463 (2004).

[31] Y. M. Rhee and M. Head-Gordon, Scaled second order perturbation corrections to configuration interaction singles: efficient and reliable excitation energy methods, J. Phys. Chem. A **111**, 5314 (2007).

[32] A. I. Krylov, Equation-of-motion coupled-cluster methods for open-shell and electronically excited species: The hitchhiker's guide to Fock space, Annu. Rev. Phys. Chem. **59**, 433 (2008).

[33] A.A. Granovsky, Extended multi-configuration quasi-degenerate perturbation theory: The new approach to multi-state multi-reference perturbation theory, J. Chem. Phys. **134**, 214113 (2011).

[34] K. B. Bravaya, M. G. Khrenova, B. L. Grigorenko, A. V. Nemukhin, and A. I. Krylov, Effect of protein environment on electronically excited and ionized states of the green fluorescent protein chromophore, J. Phys. Chem. B **115**, 8296 (2011).

[35] K. B. Bravaya, O. M. Subach, N. Korovina, V. V. Verkhusha, and A. I. Krylov, Insight into the common mechanism of the chromophore formation in the red fluorescent proteins: The elusive

blue intermediate revealed, J. Am. Chem. Soc. **134**, 2807 (2012).

[36] A. M. Bogdanov, A. Acharya, A. V. Titelmayer, A. V. Mamontova, K. B. Bravaya, A. B. Kolomeisky, K. A. Lukyanov, and A. I. Krylov, Turning on and off photoinduced electron transfer in fluorescent proteins by $\pi$-stacking, halide binding, and Tyr145 mutations, J. Am. Chem. Soc. **138**, 4807 (2016).

[37] T. Sen, A. V. Mamontova, A. V. Titelmayer, A. M. Shakhov, A. A. Astafiev, A. Acharya, K. A. Lukyanov, A. I. Krylov, and A. M. Bogdanov, Influence of the first chromophore-forming residue on photobleaching and oxidative photoconversion of EGFP and EYFP, Int. J. Mol. Sci. **20**, 5229 (2019).

[38] Shao, Y.; Gan, Z.; Epifanovsky, E.; Gilbert, A.T.B.; Wormit, M.; Kussmann, J.; Lange, A.W.; Behn, A.; Deng, J.; Feng, X., et al., Advances in molecular quantum chemistry contained in the Q-Chem 4 program package, Mol. Phys. **113**, 184 (2015).

[39] A. I. Krylov and P. M. W. Gill, Q-Chem: An engine for innovation, WIREs: Comput. Mol. Sci. **3**, 317 (2013).

[40] A. A. Granovsky, PC GAMESS/Firefly, http://classic.chem.msu.su/gran/gamess.

[41] J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R.D. Skeel, L. Kale, and K. Schulten, Scalable molecular dynamics with NAMD, J. Comput. Chem. **26**, 1781 (2005).

[42] F. Plasser, M. Wormit, and A. Dreuw, New tools for the systematic analysis and visualization of electronic excitations. I. Formalism, J. Chem. Phys. **141**, 024106 (2014).

[43] K. B. Bravaya, B. L. Grigorenko, A. V. Nemukhin, and A. I. Krylov, Quantum chemistry behind bioimaging: Insights from ab initio studies of fluorescent proteins and their chromophores, Acc. Chem. Res. **45**, 265 (2012).

[44] E. Epifanovsky, I. Polyakov, B. L. Grigorenko, A. V. Nemukhin, and A. I. Krylov, Quantum chemical benchmark studies of the electronic properties of the green fluorescent protein chromophore: I. Electronically excited and ionized states of the anionic chromophore in the gas phase, J. Chem. Theory Comput. **5**, 1895 (2009).

[45] A. V. Nemukhin, B. L. Grigorenko, M. Khrenova, and A. I. Krylov, Computational challenges in modeling of representative bioimaging proteins: GFP-like proteins, flavoproteins, and phytochromes, J. Phys. Chem. B **123**, 6133 (2019).

[46] R. B. Vegh, K. B. Bravaya, D. A. Bloch, A. S. Bommarius, L. M. Tolbert, M. Verkhovsky,

A. I. Krylov, and K. M. Solntsev, Chromophore photoreduction in red fluorescent proteins is responsible for bleaching and phototoxicity, J. Phys. Chem. B **118**, 4527 (2014).

[47] B. L. Grigorenko, A. V. Nemukhin, D. I. Morozov, I. V. Polyakov, K. B. Bravaya, and A. I. Krylov, Toward molecular-level characterization of photo-induced decarboxylation of the green fluorescent protein: Accessibility of the charge-transfer states, J. Chem. Theory Comput. **8**, 1912 (2012).

[48] B. L. Grigorenko, A. V. Nemukhin, I. V. Polyakov, M. G. Khrenova, and A. I. Krylov, A light-induced reaction with oxygen leads to chromophore decomposition and irreversible photobleaching in GFP-type proteins, J. Phys. Chem. B **119**, 5444 (2015).

[49] P. D. Dahlberg, A. M. Sartor, J. Wang, S. Saurabh, L. Shapiro, and W. E. Moerner, Identification of PAmKate as a red photoactivatable fluorescent protein for cryogenic super-resolution imaging, J. Am. Chem. Soc. **140**, 12310 (2018).

[50] P. D. Dahlberg, A. Saurabh, A. M. Sartor, J. Wang, P. G. Mitchell, W. Chiu, L. Shapiro, and W. E. Moerner, Cryogenic single-molecule fluorescence annotations for electron tomography reveal in situ organization of key proteins in Caulobacter, Proc. Nat. Acad. Sci. **117**, 13937 (2020).