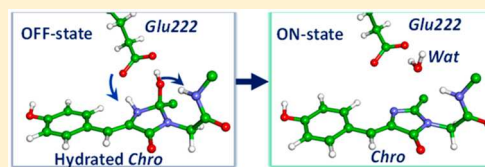


Computational Modeling Reveals the Mechanism of Fluorescent State Recovery in the Reversibly Photoswitchable Protein Dreiklang

Bella L. Grigorenko,^{†,‡} Igor V. Polyakov,^{†,‡} Anna I. Krylov,^{§,‡} and Alexander V. Nemukhin^{*,†,‡,§}[†]Department of Chemistry, Lomonosov Moscow State University, Moscow 119991, Russia[‡]Emanuel Institute of Biochemical Physics, Russian Academy of Sciences, Moscow 119334, Russia[§]Department of Chemistry, University of Southern California, Los Angeles, California 90089-0482, United States

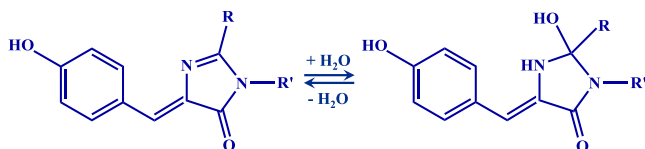
ABSTRACT: The unique properties of the photoswitchable protein Dreiklang are attributed to a reversible hydration/dehydration reaction at the imidazolinone ring of the chromophore. Recovery of the fluorescent state, which is associated with a chemical reaction of the chromophore's dehydration, is an important part of the photocycle of this protein. Here we characterize the fluorescent (ON) and nonfluorescent (OFF) states of Dreiklang and simulate the thermal recovery reaction OFF → ON using computational approaches. By using molecular modeling methods including the quantum mechanics/molecular mechanics (QM/MM) technique, we characterize the structures and spectra of the ON- and OFF-states. The results are consistent with available experimental data. The computed reaction profile explains the observed recovery reaction and clarifies the mechanism of chemical transformations in the chromophore-containing pocket in Dreiklang.



INTRODUCTION

It is difficult to overstate the significance of fluorescent proteins from the green fluorescent protein (GFP) family widely used as biomarkers in living cells.¹ A new wave of interest emerged due to the development of super-resolution microscopy.² In these applications, photoswitchable proteins play a central role. In contrast to passive fluorescent reporters, optical properties of photoswitchable proteins can be modified by light. For example, their ability to fluoresce can be switched on and off, and their colors can be changed as well. The engineered protein Dreiklang³ is one of such promising species. The unique properties of Dreiklang are attributed to a reversible hydration/dehydration reaction at the imidazolinone ring of the chromophore, as summarized in Scheme 1. It

Scheme 1. Reversible Hydration/Dehydration Reaction of the Dreiklang Chromophore



is a rare example of a reversible photochemical modification of the GFP-like chromophore inside the chromophore-containing pocket. Another unusual aspect of this transformation is that, upon illumination, a water molecule in the chromophore-containing pocket reacts with the chromophore and that the original chromophore is restored via the dehydration reaction.

Dehydration reactions of the GFP-like chromophore are believed to be a part of the chromophore maturation, a post-translational modification of the Ser65-Tyr66-Gly67 tripep-

tide.^{4,5} However, much less is known about the reversible photochemical hydration of the chromophore in Dreiklang. The mechanistic proposal (Scheme 1) was put forward in the original paper.³ A recent study⁶ reported the results of femtosecond experiments with Dreiklang, tentatively suggesting additional details of the mechanism of the hydration reaction. No proposals for the dehydration reaction, i.e., the recovery of the fluorescent state, have been formulated so far.

The structure of Dreiklang is typical for the GFP-like proteins. The chromophore (Chro) resides in an α -helical segment, enclosed by a β -barrel. Unlike the parent GFP variants in which chromophores are autocatalytically formed from the Ser65(Thr65)-Tyr66-Gly67 tripeptide, the chromophore in Dreiklang is formed by the Gly65-Tyr66-Gly67 residues. The neighboring side chains at positions 64, 68, and 203 also differ: Phe64, Val68, and Thr203 in GFP, versus Ile64, Leu68, and Tyr203 in Dreiklang. Upon the chromophore's maturation, Dreiklang exists in the fluorescent ON-state, which can be converted to the OFF-state illumination by 3.06 eV (405 nm) light.³

The crystal structure reveals³ that the chromophore in the OFF-state is hydrated at the imidazolinone ring. The recovery of the ON-state occurs either photochemically, by illumination at 3.39 eV (365 nm), or in the dark. In the latter case, the ON-state recovers spontaneously with a half-life of ~ 10 min at room temperature.³ Using the Eyring–Polanyi equation and assuming a transmission frequency of 1, these kinetic data correspond to an OFF → ON Gibbs free energy barrier of 21 kcal/mol.

Received: July 23, 2019

Revised: September 24, 2019

Published: October 1, 2019

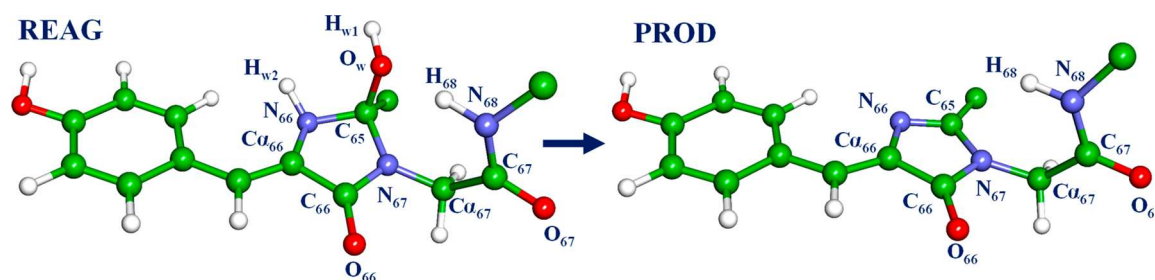


Figure 1. Structures of the hydrated (left) and original (right) chromophore in Dreiklang. Here and in other figures, carbon atoms are shown in green, oxygen in red, nitrogen in blue, and hydrogen in light gray.

The unique photoswitchable properties of Dreiklang are attributed to the chemical transformations of the chromophore. Detailed mechanistic understanding of the photo-conversion is important for developing guidelines for the rational design of reversibly switchable fluorescent labels. Moreover, the significance of the dehydration reaction at the imidazolinone ring in Dreiklang extends beyond this particular protein, as the reactions of this type occur upon maturation of the GFP-like chromophore.^{7–9}

In this paper, we use methods of computational chemistry to model the recovery reaction (OFF $\rightarrow$ ON) of the fluorescent state of Dreiklang. Successful applications of computational chemistry to modeling properties of photoactive proteins are summarized in recent review papers;^{10–13} these studies include simulations of reversibly switchable fluorescent proteins¹⁴ as well as investigations of photochemical and thermal reaction pathways in photoreceptors.^{15,16}

Figure 1 illustrates the reaction in terms of molecular models of the chromophore in the hydrated and dehydrated forms. In Figure 1, we also show designation of the key atoms of reacting species. Because here we model the reaction of the ON-state recovery, we place the hydrated species (the chromophore in the OFF-state) in the left and denote molecular compositions in the OFF- and ON-states as REAG (reagents) and PROD (products), respectively.

A viable reaction mechanism should explain the cleavage of the $C_{65}-O_w$ bond, the formation of the water molecule, and the restoration of the original chromophore. As we show below, a mechanism consistent with the observed relatively low energy barrier in the thermal reaction involves the participation of the backbone chain $O_{67}-C_{67}-N_{68}-H_{68}$ and the Glu222 side chain, which is located unusually close to the chromophore's imidazolinone ring.

MODELS AND METHODS

To construct model systems representing the protein-bound chromophore, we started from the coordinates of heavy atoms in the crystal structure PDB ID 3ST2³ (chain A) from the Protein Data Bank.¹⁷ This structure, solved with 1.9 Å resolution, corresponds to the initially formed protein crystal of Dreiklang.³ This fluorescent equilibrium state was then switched to the nonfluorescent OFF-state by irradiation at 405 nm until fluorescence reached a minimum; the corresponding crystal structure PDB ID 3ST3 was solved with 1.7 Å resolution.³ Then the ON-state was restored by irradiation with 365 nm light until the fluorescence reached a maximum; the corresponding crystal structure PDB ID 3ST4 was solved with 2.0 Å resolution.³

In simulations, we added hydrogen atoms using molecular mechanics tools; we assumed that side chains of Arg and Lys

are positively charged and that side chains of Glu and Asp are negatively charged. The protein was fully surrounded by explicit water molecules. The system with the original (nonhydrated) chromophore was equilibrated in preliminary molecular dynamics simulations with the CHARMM force field¹⁸ and the chromophore's parameters from ref 19.

The choice of the QM subsystem for the quantum mechanics/molecular mechanics (QM/MM) calculations^{20,21} was based on the previous studies of the GFP-like proteins.^{9,13,22,23} A large fraction of the chromophore-containing pocket was assigned to the QM part. Specifically, atoms of Chro shown in the left side in Figure 1, side chains of Gln94, Arg96, His145, Tyr203, Ser205, and Glu222, and seven water molecules were included in QM. Calculations of energies and forces in QM were carried out using Kohn–Sham DFT with the M06-L functional²⁴ and the cc-pVDZ basis set. The AMBER force field was used in MM. The NWChem software package²⁵ was used to scan fragments of potential energy surface along the assumed reaction coordinates. Unconstrained QM/MM minimization allowed us to locate minimum energy points corresponding to the REAG, PROD, and INT1–INT3 structures. Transition states were first located in series of constrained QM/MM minimizations and then confirmed by frequency calculations. Gradual descent in the forward and backward directions from the TS points verified that the true saddle points were found.

Vertical excitation energies $S_0 \rightarrow S_1$ at selected points on the ground-state potential energy surface were computed using XMCQDPT2²⁶ with the cc-pVDZ basis set, the protocol that we verified earlier and used extensively in studies of the photoreceptor proteins.¹³ Here, the perturbation theory calculations were based on the CASSCF wave functions obtained by distributing 16 electrons over 12 orbitals and using density averaging over 11 states. The active space included orbitals from Chro and Thr203, which interact via π -stacking.

RESULTS

Computational Characterization of the ON- and OFF-States. The main focus of this paper is on the mechanism of the dehydration reaction connecting the OFF- and ON-states in Dreiklang. In this subsection, we validate our models against the available experimental data.^{3,6} The original experimental study³ has shown that the protein in the equilibrium state (PDB ID 3ST2) is converted to the inactive OFF-state by illumination by 365 nm light (3.40 eV). The corresponding crystal structure is PDB ID 3ST3. Subsequent illumination of the OFF-state by 405 nm (3.06 eV) light results in the fluorescent state ON; the corresponding crystal structure is PDB ID 3ST4. In our simulations, we started from the coordinates of heavy atoms of the PDB ID 3ST2 entry and

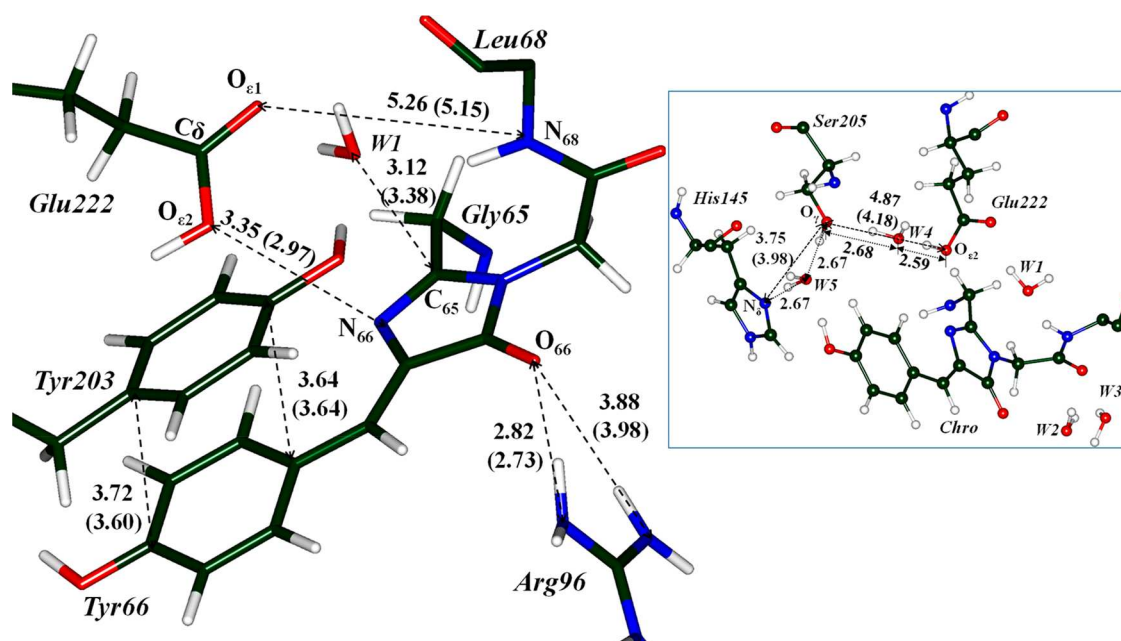


Figure 2. Fragment of the chromophore-containing pocket in the ON-state in Dreiklang. QM/MM-optimized distances between heavy atoms are in Å; the values in parentheses are from the crystal structure PDB ID 3ST4 (chain A). The inset illustrates the system from another perspective showing a proton wire $O_{\epsilon 2}(\text{Glu222})-W4-O_{\gamma}(\text{Ser205})-W5-N_{\delta}(\text{His145})$.

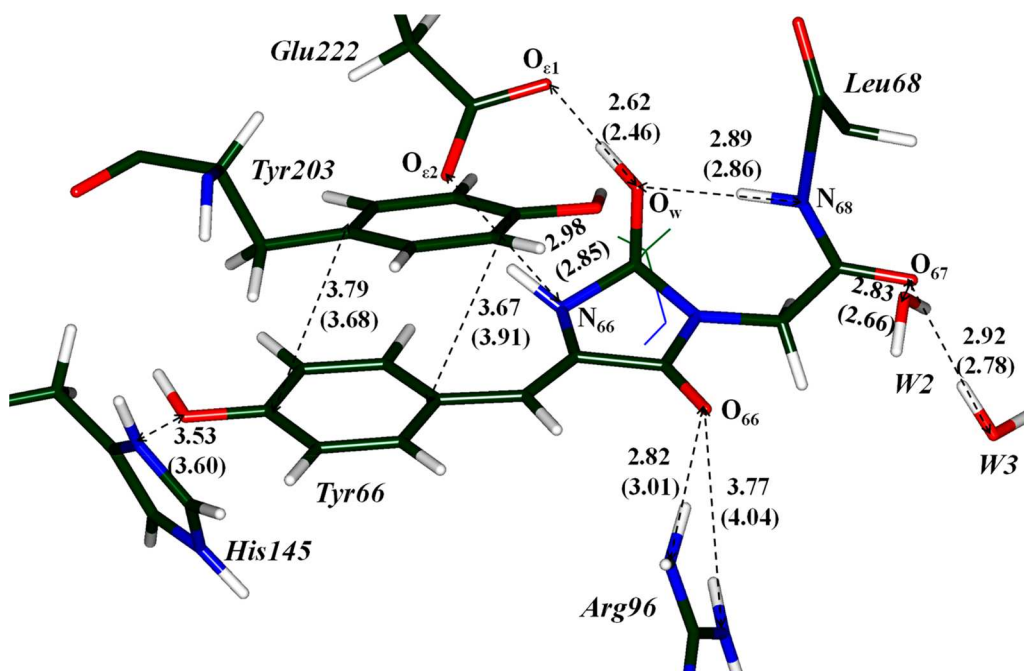


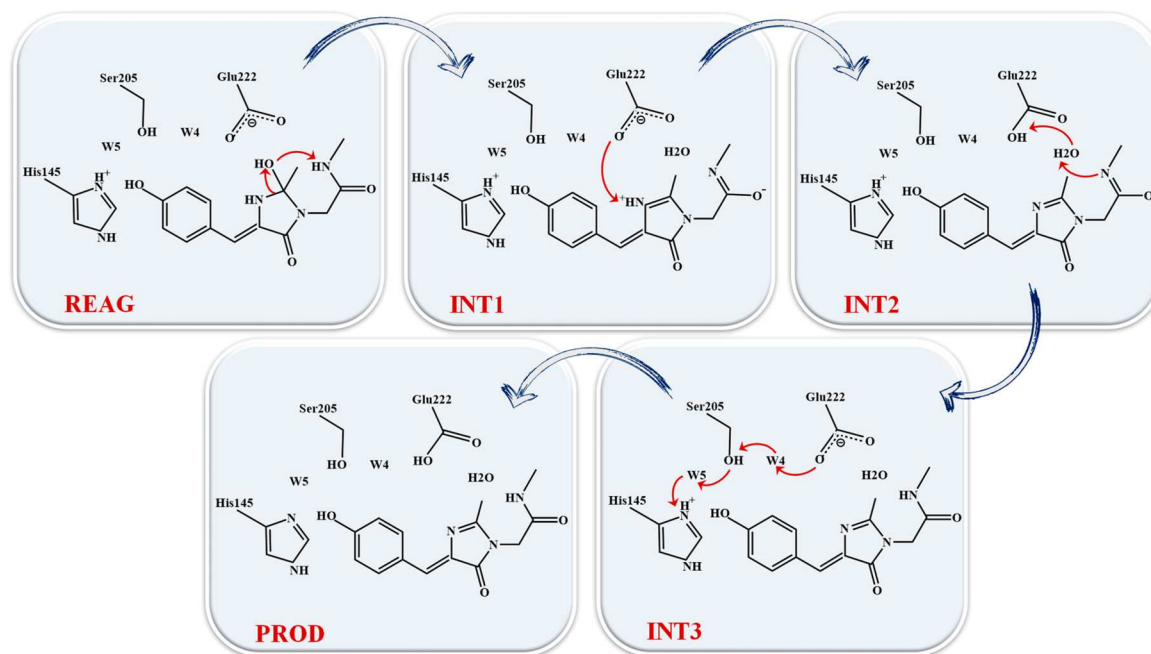
Figure 3. Fragment of the chromophore-containing pocket in the OFF-state in Dreiklang. The QM/MM-optimized distances between heavy atoms are in Å; values in parentheses refer to those from the crystal structure PDB ID 3ST3 (chain A).

prepared an all-atom protein model as described in [Models and Methods](#). The QM/MM optimization produced a model system with the neutral chromophore, neutral His145, and neutral (protonated) Glu222, representing the system assumed to be the ON-structure in the experimental paper.⁶ Its optimized geometry should be compared to the crystal structure PDB ID 3ST4, and the excitation energy should be compared with the absorption band maximum at 3.01 eV (411 nm). A model system with the hydrated chromophore ([Figure 1](#), left) was prepared manually from our ON-model. Its

structure should be compared with the crystal structure PDB ID 3ST3 (the OFF-state), and its excitation energy should be compared with the absorption band maximum at 3.65 eV (339 nm).

[Figure 2](#) shows a fragment of the chromophore-containing pocket in the ON-state. We compare its key structural parameters with the parameters of the chain A from the entry PDB ID 3ST4. The water molecule W1 in the model structure ([Figure 2](#)) is located in the cavity sufficiently close to the carbon atom C₆₅ (3.12 Å in the model, 3.38 Å in the

Scheme 2. Intermediates of the Reaction of Thermal Recovery of the ON-State



crystal) to form the hydrated chromophore moiety. While simulations of the photoinduced chemical reaction of formation of the hydrated chromophore are beyond the scope of this study, we simply point out that our model system for the ON-state (the neutral Chro, deprotonated His145, and protonated Glu222) does not contradict the early mechanistic hypotheses.⁶

We note the proximity of the Glu222 side chain to the imidazolinone ring and to the group of backbone atoms ($H_{68}-N_{68}-C_{67}-O_{67}$) connecting the Gly67 (partly converted to the chromophore moiety) and Leu68 side chains; the relevant distances are between the $O_{e1}(Glu222)-N_{68}$ atoms (5.26 Å in the model, 5.15 Å in the crystal) and between the $O_{e2}(Glu222)-N_{66}$ atoms (3.35 Å in the model, 2.97 Å in the crystal). The positions of the chromophore groups in the model system relative the important residues Tyr203 and Arg96 also agree well with the crystal structure.

The inset in Figure 2 shows the model for the ON-state from another perspective. We draw the attention to a perfect proton wire $O_{e2}(Glu222)-W4-O_{\gamma}(Ser205)-W5-N_{\delta}(His145)$ connecting the Glu222 and His145 side chains. The hydrogen-bond chain $O_{e2}(Glu222)-O_w(W4)-O_{\gamma}(Ser205)-O_w(W5)-N_{\delta}(His145)$ is characterized by distances between the neighboring heavy atoms 2.6–2.7 Å. It guarantees an efficient proton transfer along this proton wire.^{27–29} In GFP, a similar proton wire connects the phenolic oxygen of Chro with the $O_e(Glu222)$ via Ser205 and a water molecule.²² Proton transfer along such a route is an essential part of the GFP machinery.³⁰ In Dreiklang, the corresponding proton wire should also be operable; the computed potential energy profile shows that the structure with the protonated Glu222 and unprotonated His145 is about 5 kcal/mol lower in energy than that with the unprotonated Glu222 and protonated His145. The computed energy of the lowest vertical transition with a large oscillator strength is 2.90 eV (obtained with XMCQDPT2). The corresponding wavelength of 427 nm agrees well with the experimental absorption band maximum at 411–413 nm.

We constructed a model for the OFF-state on the basis of the computationally derived ON-state model. We split the water molecule W1 (see Figure 2) to the O_wH_{w1} and H_{w2} moieties, which were attached to the C_{65} and N_{66} atoms in the ring; this corresponds to proposed nucleophilic water addition across the $C=N$ bond of the imidazolinone ring. We then carried out QM/MM optimizations considering different protonation states of the chromophore, His145 and Glu222. We concluded that the structure with the neutral Chro, anionic Glu222, and protonated His145 is the most likely candidate for representing the OFF-state (a detailed report will be published elsewhere).

Figure 3 shows the chromophore-containing pocket in the OFF-state (which corresponds to the REAG structure in the dehydration reaction). We compare the computed distances between heavy atoms from different molecular groups with those in the chain A of the crystal structure PDB ID 3ST3 (values in parentheses). In the context of dehydration reaction, the following short distances are of primary importance: between the backbone atom N_{68} and the oxygen atom O_w , 2.89 Å in the model, 2.86 Å in the crystal; between O_w and $O_{e1}(Glu222)$, 2.62 Å in the model, 2.46 Å in the crystal; between N_{66} and $O_{e2}(Glu222)$, 2.98 Å in the model, 2.85 Å in the crystal. The latter two distances clearly show that the side chain Glu222 is deprotonated in the OFF-state. As shown below, proton transfers along the corresponding pathways account for the reaction. Also, two water molecules (called W2 and W3 in Figure 3) hydrogen-bonded to O_{67} play an important role in the reaction, serving as the oxyanion hole. In the model structure, these two water correspond to crystal waters #5 and #12 (chain A in PDB ID 3ST3).

The computed energy for vertical transition $S_0 \rightarrow S_1$ with a large oscillator strength is 3.66 eV (obtained with XMCQDPT2). This value, corresponding to the wavelength 338 nm, agrees well with the experimental absorption band maximum at 339 nm. To sum up, a good agreement of computed structural parameters and the absorption band

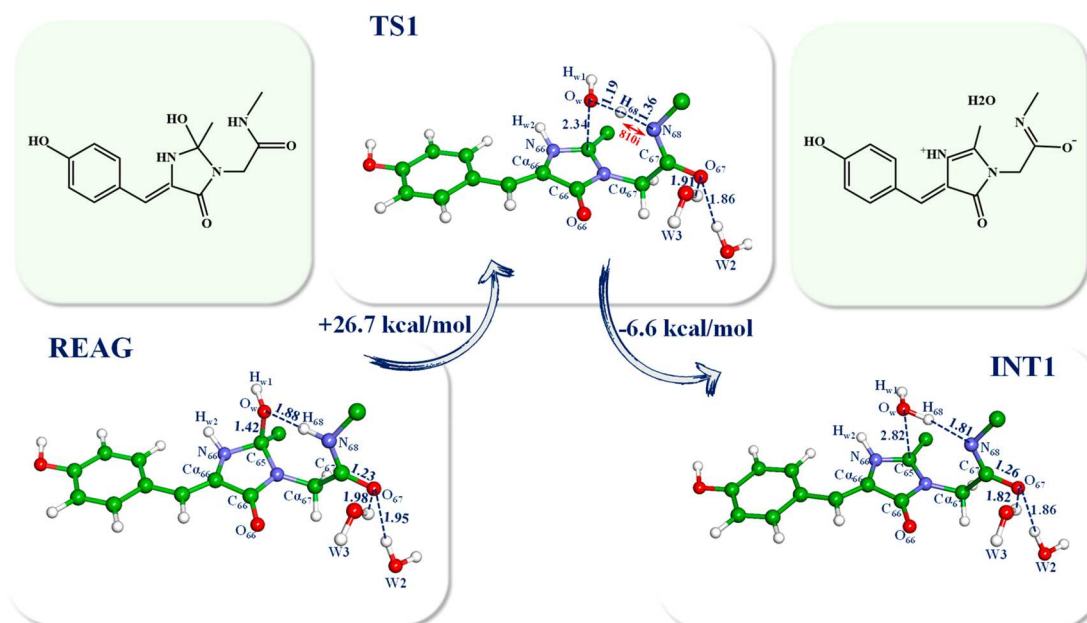


Figure 4. Molecular models of the reaction species at the elementary step REAG $\rightarrow$ TS1 $\rightarrow$ INT1. Distances are in Å.

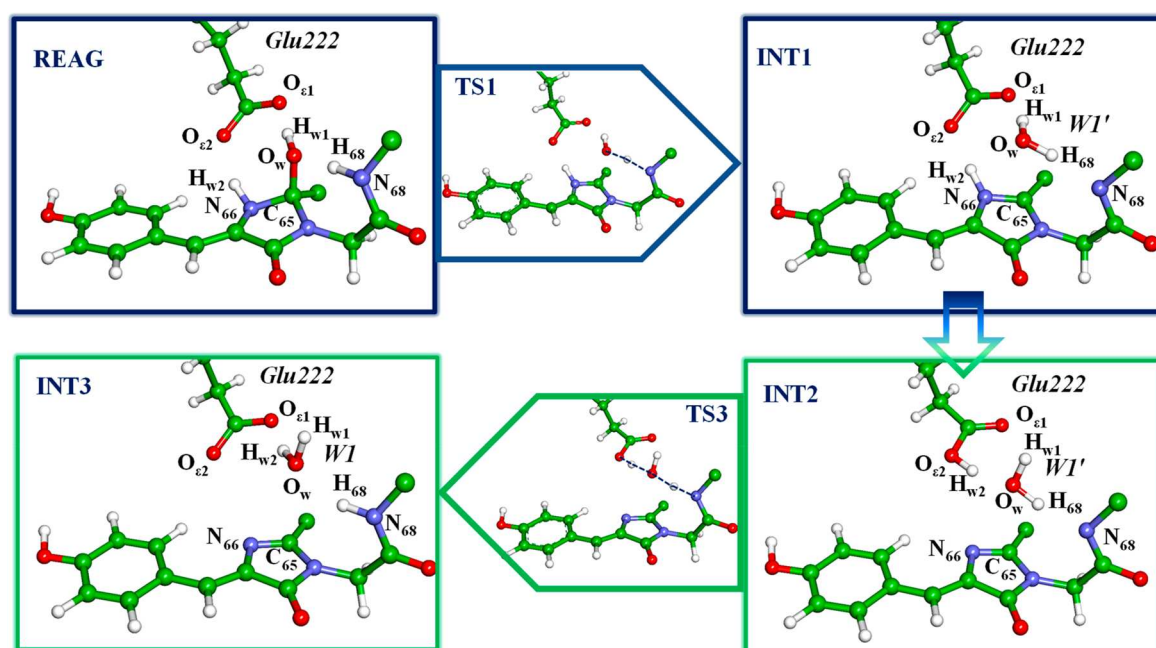


Figure 5. Intermediates and transition states of the dehydration reaction illustrating the role of Glu222.

maxima with the corresponding experimental data provide a validation of our model systems for the ON- and OFF-states.

Pathway of Thermal Recovery Reaction. We computed a minimum energy profile for the reaction of the thermal recovery of the ON-state in a series of constrained QM/MM minimizations by considering relevant reaction coordinates at each elementary step.

Scheme 2 illustrates the reaction pathway from REAG (OFF-state) to PROD (ON-state) on the ground electronic state potential energy surface.

Calculations reveal that the rate-determining step (REAG $\rightarrow$ INT1) in this reaction is the O_w-H_{w1} protonation by the backbone amide $N_{68}-H_{68}$ (see Figure 1, left), resulting in the formation of water. The transition from INT1 to INT2

involves transient protonation of Glu222, and the next step is a proton transfer in the chain $O_{e2}(\text{Glu222})-W1'-N_{68}$, restoring the structure of the nonhydrated chromophore (see Figure 1, right). The intermediate INT3 corresponds to a local minimum on the potential energy surface, while the transition to the global minimum (ON-state) requires concerted proton transfer along the $O_{e2}(\text{Glu222})-W4-O_{\gamma}(\text{Ser205})-W5-N_{\delta}(\text{His145})$ wire. At all elementary steps beyond the first one, the Glu222 side chain serves as a proton shuttle.

In the following, we consider in details the rate-determining step of the thermal reaction OFF $\rightarrow$ ON in Dreiklang and then describe the steps of proton shuttling via Glu222.

Rate-Determining Step. The elementary step with the highest barrier on the route from REAG to PROD is the

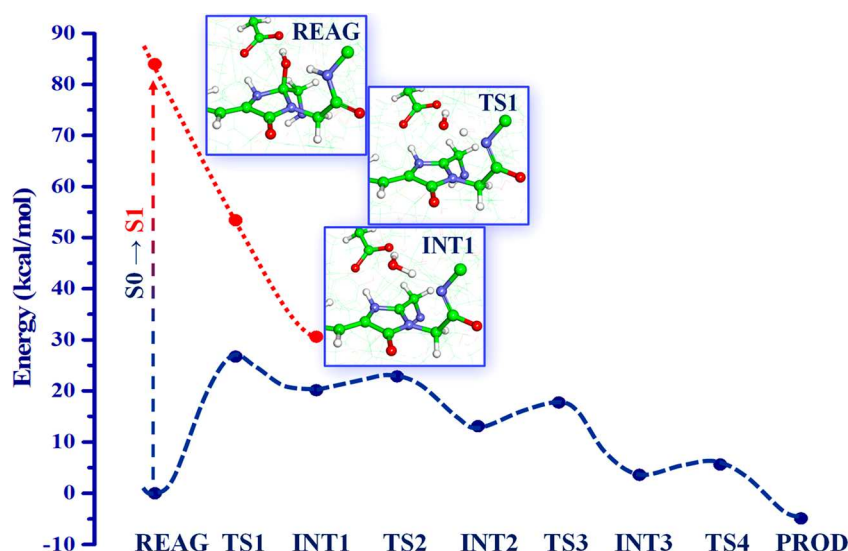


Figure 6. Energetics of the ON-state recovery reaction in Dreiklang. Panels show the structures where estimates of the excited state energies were performed.

cleavage of the $C_{65}-O_w$ bond in the hydrated chromophore and the formation of a water molecule from the O_w-H_w and H_{68} moieties. As shown in Figure 4, the hydrogen-bonded pattern $N_{68}-H_{68}-O_w$ is perfectly aligned for the proton transfer. This elementary step leads to the reaction intermediate INT1 separated from REAG by the transition state TS1. A short initial distance between O_w and H_{68} , 1.88 Å, is due to the proximity of the Leu68 side chain to the chromophore (Figure 2), which, in turn, can be explained by the small size of the residue in position 65: Gly65 in Dreiklang versus Ser or Thr in regular GFPs.

The preliminary scan of the ground-state PES along the two coordinates corresponding to the $C_{65}-O_w$ and $N_{68}-H_{68}$ distances allowed us to locate an intermediate conformation, starting from which the single reaction coordinate corresponding to the $N_{68}-H_{68}$ stretch could be considered. We located the saddle point corresponding to the TS1 structure, from which the unconstrained descents in the forward and backward directions led to the REAG and INT1 points on the PES. Figure 4 shows molecular models of the corresponding structures. The produced water molecule $W1'$ bridges the carboxyl group of Glu222 and the backbone fragment.

Frequency calculations performed for these three stationary points (REAG, TS1, INT1) confirmed that REAG and INT1 are the local minima and the TS1 structure has a single imaginary frequency $810i\text{ cm}^{-1}$. The mode of this imaginary frequency corresponds to the $N_{68}-H_{68}$ stretch. The electronic energy difference between the stationary points TS1 and REAG on the QM/MM potential energy surface is 27 kcal/mol. Due to the high imaginary frequency, zero-point energy corrections considerably reduce the barrier height, making it closer to 21 kcal/mol. We conclude that theoretical and experimental data on the kinetics of the thermal reaction of dehydration in Dreiklang are in agreement, taking into account typical errors in DFT calculations of barrier heights (up to several kcal/mol).³¹

The reaction intermediate INT1 lies about 20 kcal/mol above the OFF-structure. As shown in Figure 4, the changes in the geometry parameters upon the route from REAG to INT1 are consistent with the changes in the electronic structure: there is a decrease in the $C_{67}-N_{68}$ bond length and increase in

the $C_{67}-O_{67}$ bond length. The developing negative charge at the $C_{67}-O_{67}$ group (+0.62 on C_{67} and -0.71 on O_{67}) is stabilized by the two approaching water molecules, $W2$ and $W3$, which play a role of the oxyanion hole. The positive charge +0.49 refers to the proton H_{w2} attached to the N_{66} atom.

Glu222 as a Proton Shuttle in the Dehydration Reaction. Figure 5 illustrates the dehydration reaction by showing the computed structures along the formation of the water molecule from the O_w-H_w and H_w moieties initially bound to the C_{65} and N_{66} atoms of the hydrated chromophore in REAG. Figure 5 and Scheme 2 propose a reaction route with a reasonable energetics connecting REAG (left bottom panel in Figure 5) and INT3 (right bottom panel in Figure 5), which differ by the hydrated/dehydrated imidazolinone ring. The side chain of Glu222 and the backbone atoms $O_{67}-C_{67}-N_{68}-H_{68}$ participate in the intermediate transformations only.

The step REAG $\rightarrow$ INT1 was described in the preceding subsection (Figure 4). Subsequent step (INT1 $\rightarrow$ TS2 $\rightarrow$ INT2) involves the transfer of proton H_{w2} from N_{66} to O_{e2} (Glu222) along the corresponding hydrogen bond. The reaction coordinate at this step is the $N_{66}-H_{w2}$ distance; the energy of the TS2 structure is 2.7 kcal/mol above INT1.

The structure of INT2 (central panel in Figure 5) shows a perfect hydrogen bond pattern O_{e2} (Glu222)- $H_{w2}-O_w-H_{68}-N_{68}$. Bottom right inset in Figure 5 illustrates a concerted proton transfer along this route via the transition state TS3. The energy of TS3 is 4.6 kcal/mol above INT2.

Although the structure of INT3 corresponds to the product of the dehydration reaction, its energy is slightly (3.6 kcal/mol) above the REAG structure. This is because INT3 is not the lowest energy configuration of Dreiklang in the ON-state. The inset in Figure 2 shows that structures with the same chromophore, but with different protonation states of Glu222 and His145 are connected by a perfect proton wire O_{e2} (Glu222)- $W4-O_\gamma$ (Ser205)- $W5-N_\delta$ (His145). Consequently, a lower energy stationary point PROD (see Figure 2) is reached upon the concerted proton transfer along the wire. The corresponding transition state TS4 is 5 kcal/mol above INT3. This type of easy proton translocation along proton wires including the side chain of Glu222 is a well-

known feature of GFP.^{22,27–29,32} In general, proton transfer along perfectly aligned hydrogen-bond patterns occurs with low activation energies.^{33,34}

Energy Profiles. The computed ground-state energies of the key structures are summarized in black in Figure 6. The point PROD is below REAG by about 5 kcal/mol, consistently with a spontaneous decay of the OFF state in the dark. Low activation energies (less than 5 kcal/mol) of all elementary steps beyond the first one can be explained by the nature of these transformations (proton transfer along perfectly aligned hydrogen-bonded residues).

Although here we do not consider excited-state reactivity, we draw the attention to some details relevant to the light-induced recovery of the Dreiklang's ON-state. We remind the reader that the computed $S_0 \rightarrow S_1$ vertical excitation energy of the OFF-state (3.66 eV, ~ 85 kcal/mol or 338 nm), matches perfectly the observed absorption band maximum at 340 nm. Experimentally, the OFF $\rightarrow$ ON switching is induced by 405 nm light, which corresponds to 3.05 eV or 70 kcal/mol. Using the XMCQDPT2 method, we computed the vertical excitation energies $S_0 \rightarrow S_1$ for the REAG, TS1, and INT1 structures. Figure 6 shows the obtained points in red; they are connected by the dotted lines to guide the eye. Importantly, these excited-state energies show a gradual decrease toward the ground-state energy until the point INT1. These estimates strongly suggest that the observed light-induced recovery of the ON-state proceeds along the critical segment REAG $\rightarrow$ TS1 $\rightarrow$ INT1 closely to that shown in red in Figure 6, and the conical intersection point S_1/S_0 is likely to occur in the vicinity of INT1. Once INT1 is reached, further steps are the same as in the thermal mechanism proposed in this work.

DISCUSSION

The mechanism of the chromophore's dehydration reaction revealed by the simulations presents an interesting and rare example of chemical transformations in proteins involving the participation of the backbone chain. We point out a recent study³⁵ of enzymatic process describing backbone N-methylation in peptides. A specific reacting complex facilitates the removal of the amide proton and stabilizes the resulting negative charge activating the amide bond for methylation. A similar process takes place in the dehydration reaction in Dreiklang at the step REAG $\rightarrow$ INT1 (Figure 4). A highly perturbed structure of the imidazolinone ring and the proximity of the hydroxyl O_{wH_w} to the backbone in REAG facilitate the removal of the amide proton H_{68} , while the developing negative charge on O_{67} is stabilized by the oxyanion hole formed by two water molecules (W2 and W3). The subsequent step (Figure 5) leads to the intermediate INT3 with the restored structure of the backbone chain and the GFP-like chromophore. The proton shuttle Glu222 is the key player at this reaction segment.

Thus, a critically important aspect of the switching properties of Dreiklang and the hydration/dehydration reaction is the proximity of the Glu222 side chain to the imidazolinone ring and to the backbone atoms (H_{68} – N_{68} – C_{67} – O_{67}) connecting the Gly67 (partly converted to the chromophore moiety) and Leu68 side chains. The original paper³ states that the introduction of Gly65 and Tyr203 instead of Ser(Thr)65 and Thr203 in the parent GFP species accounts for the short distance between Glu222 and the imidazolinone ring. The comparison of the computed model systems in Dreiklang (main body in Figure 2) with wt-GFP²²

illustrates this observation; for example, the distance between the carbon atoms C_δ in Glu222 and C_{65} in the ring is considerably shorter (by almost 1 Å) in Dreiklang. The same holds for the distance between the atom $O\epsilon 1$ in Glu222 and the carbonyl oxygen in the side chain at position 68 (Leu in Dreiklang or Val in GFP).

Our simulations also highlight the role of proton wires, i.e., the patterns of aligned hydrogen-bonded residues along which the proton can be translocated with relatively low activation energy. In Dreiklang, this is illustrated in the inset in Figure 2; the corresponding pathway connects the INT3 and PROD structures. Proton wires are important for understanding various properties of proteins including GFP.^{22,27–29}

In this work, we do not discuss photochemical reactions in Dreiklang beyond a tentative mechanistic suggestion for the light-induced recovery reaction (Figure 6). A careful study of photoinduced reactions requires expensive calculations of excited-state dynamics in model systems. We hope that the structures described in this work will be helpful in such simulations; the work along these lines is in progress.

CONCLUSION

The recovery reaction of the fluorescent state, OFF $\rightarrow$ ON transformation in Dreiklang, is an important part of its reversible photoswitching cycle, which involves hydration/dehydration of the chromophore. The mechanism of this reaction in the ground electronic state revealed by molecular modeling is consistent with the structural features of the chromophore-containing pocket in this protein from the GFP family. Specifically, a smaller size of the residue Gly at position 65 relative to Ser or Thr in other GFPs is responsible for the proximity of the chromophore's imidazolinone ring to the critical residue Glu222 and to a fragment of the backbone chain O_{67} – C_{67} – N_{68} – H_{68} . This motif facilitates the dehydration reaction with a relatively low energy barrier in the rate-determining step, in agreement with the observed recovery reaction rate in the dark. The established reaction route connects the OFF- and ON-states of the protein. The computed absorption spectra of the relevant structures are consistent with the available experimental data. This work and related studies of the reactions in chromophore-containing pockets in GFP-like proteins demonstrate a diversity of mechanisms of chemical transformations occurring in these fascinating systems.

AUTHOR INFORMATION

Corresponding Author

*Mailing address: Chemistry Department, Lomonosov Moscow State University, Leninskie Gory 1/3, Moscow, 119991, Russian Federation. E-mail: anemukhin@yahoo.com; anem@lcc.chem.msu.ru.

ORCID

Anna I. Krylov: 0000-0001-6788-5016

Alexander V. Nemukhin: 0000-0002-4992-6029

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by the Russian Science Foundation (Project #17-13-01051) to B.L.G. and I.V.P. and by the U.S. National Science Foundation (No. CHE-1856342) to A.I.K. A.I.K. is also a grateful recipient of the Simons Fellowship in

Theoretical Physics and Mildred Dresselhaus Award from CFEL/DESY, which supported her sabbatical stay in Germany. The research is carried out using the equipment of the shared research facilities of HPC computing resources at Lomonosov Moscow State University supported by the project RFME-FI62117X0011. The use of supercomputer resources of the Joint Supercomputer Center of the Russian Academy of Sciences is also acknowledged.

REFERENCES

- (1) Rodriguez, E. A.; Campbell, R. E.; Lin, J. Y.; Lin, M. Z.; Miyawaki, A.; Palmer, A. E.; Shu, X.; Zhang, J.; Tsien, R. Y. The Growing and Glowing Toolbox of Fluorescent and Photoactive Proteins. *Trends Biochem. Sci.* **2017**, *42* (2), 111–129.
- (2) Nienhaus, K.; Ulrich Nienhaus, G. Fluorescent Proteins for Live-Cell Imaging with Super-Resolution. *Chem. Soc. Rev.* **2014**, *43* (4), 1088–1106.
- (3) Brakemann, T.; Stiel, A. C.; Weber, G.; Andresen, M.; Testa, I.; Grotjohann, T.; Leutenegger, M.; Plessmann, U.; Urlaub, H.; Eggeling, C.; et al. A Reversibly Photoswitchable GFP-Like Protein with Fluorescence Excitation Decoupled from Switching. *Nat. Biotechnol.* **2011**, *29*, 942–947.
- (4) Zimmer, M. Green Fluorescent Protein (GFP): Applications, Structure, and Related Photophysical Behavior. *Chem. Rev.* **2002**, *102* (3), 759–782.
- (5) Craggs, T. D. Green Fluorescent Protein: Structure, Folding and Chromophore Maturation. *Chem. Soc. Rev.* **2009**, *38* (10), 2865–2875.
- (6) Lacombat, F.; Plaza, P.; Plamont, M.-A.; Espagne, A. 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.* **2017**, *8*, 1489–1495.
- (7) Barondeau, D. P.; Kassmann, C. J.; Tainer, J. A.; Getzoff, E. D. Understanding GFP Chromophore Biosynthesis: Controlling Backbone Cyclization and Modifying Post-Translational Chemistry. *Biochemistry* **2005**, *44* (6), 1960–1970.
- (8) Wachter, R. M. Chromogenic Cross-Link Formation in Green Fluorescent Protein. *Acc. Chem. Res.* **2007**, *40* (2), 120–127.
- (9) Grigorenko, B. L.; Krylov, A. I.; Nemukhin, A. V. Molecular Modeling Clarifies the Mechanism of Chromophore Maturation in the Green Fluorescent Protein. *J. Am. Chem. Soc.* **2017**, *139* (30), 10239–10249.
- (10) Bravaya, K. B.; Grigorenko, B. L.; Nemukhin, A. V.; Krylov, A. I. Quantum Chemistry Behind Bioimaging: Insights from Ab Initio Studies of Fluorescent Proteins and Their Chromophores. *Acc. Chem. Res.* **2012**, *45* (2), 265–275.
- (11) Acharya, A.; Bogdanov, A. M.; Grigorenko, B. L.; Bravaya, K. B.; Nemukhin, A. V.; Lukyanov, K. A.; Krylov, A. I. Photoinduced Chemistry in Fluorescent Proteins: Curse or Blessing? *Chem. Rev.* **2017**, *117* (2), 758–795.
- (12) Gozem, S.; Luk, H. L.; Schapiro, I.; Olivucci, M. Theory and Simulation of the Ultrafast Double-Bond Isomerization of Biological Chromophores. *Chem. Rev.* **2017**, *117* (22), 13502–13565.
- (13) Nemukhin, A. V.; Grigorenko, B. L.; Khrenova, M. G.; Krylov, A. I. Computational Challenges in Modeling of Representative Bioimaging Proteins: GFP-Like Proteins, Flavoproteins, and Phytochromes. *J. Phys. Chem. B* **2019**, *123* (29), 6133–6149.
- (14) Smyrnova, D.; del Carmen Marín, M.; Olivucci, M.; Ceulemans, A. Systematic Excited State Studies of Reversibly Switchable Fluorescent Proteins. *J. Chem. Theory Comput.* **2018**, *14* (6), 3163–3172.
- (15) Khrenova, M. G.; Bochenkova, A. V.; Nemukhin, A. V. Modeling Reaction Routes from Rhodopsin to Bathorhodopsin. *Proteins: Struct., Funct., Bioinf.* **2010**, *78* (3), 614–622.
- (16) Mironov, V. A.; Khrenova, M. G.; Grigorenko, B. L.; Savitsky, A. P.; Nemukhin, A. V. Thermal Isomerization of the Chromoprotein asFP595 and its Kindling Mutant A143G: QM/MM Molecular Dynamics Simulations. *J. Phys. Chem. B* **2013**, *117* (43), 13507–13514.
- (17) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28* (1), 235–242.
- (18) MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; et al. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B* **1998**, *102* (18), 3586–3616.
- (19) Reuter, N.; Lin, H.; Thiel, W. 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* **2002**, *106* (24), 6310–6321.
- (20) Warshel, A.; Levitt, M. Theoretical Studies of Enzymic Reactions: Dielectric, Electrostatic and Steric Stabilization of the Carbonium Ion in the Reaction of Lysozyme. *J. Mol. Biol.* **1976**, *103* (2), 227–249.
- (21) Senn, H. M.; Thiel, W. QM/MM Methods for Biomolecular Systems. *Angew. Chem., Int. Ed.* **2009**, *48* (7), 1198–1229.
- (22) Grigorenko, B. L.; Nemukhin, A. V.; Polyakov, I. V.; Morozov, D. I.; Krylov, A. I. First-Principles Characterization of the Energy Landscape and Optical Spectra of Green Fluorescent Protein along the A→I→B Proton Transfer Route. *J. Am. Chem. Soc.* **2013**, *135* (31), 11541–11549.
- (23) Grigorenko, B. L.; Kots, E. D.; Krylov, A. I.; Nemukhin, A. V. Modeling of the Glycine Tripeptide Cyclization in the Ser65Gly/Tyr66Gly Mutant of the Green Fluorescent Protein. *Mendeleev Commun.* **2019**, *29*, 187–189.
- (24) Zhao, Y.; Truhlar, D. G. The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing of Four M06-Class Functionals and 12 other Functionals. *Theor. Chem. Acc.* **2008**, *120*, 215–241.
- (25) Valiev, M.; Bylaska, E. J.; Govind, N.; Kowalski, K.; Straatsma, T. P.; Van Dam, H. J. J.; Wang, D.; Nieplocha, J.; Apra, E.; Windus, T. L.; et al. NWChem: A Comprehensive and Scalable Open-Source Solution for Large Scale Molecular Simulations. *Comput. Phys. Commun.* **2010**, *181* (9), 1477–1489.
- (26) Granovsky, A. A. Extended Multi-Configuration Quasi-Degenerate Perturbation Theory: The New Approach to Multi-State Multi-Reference Perturbation Theory. *J. Chem. Phys.* **2011**, *134* (21), 214113.
- (27) Nadal-Ferret, M.; Gelabert, R.; Moreno, M.; Lluch, J. M. Transient Low-Barrier Hydrogen Bond in the Photoactive State of Green Fluorescent Protein. *Phys. Chem. Chem. Phys.* **2015**, *17* (46), 30876–30888.
- (28) Shinobu, A.; Palm, G. J.; Schierbeek, A. J.; Agmon, N. Visualizing Proton Antenna in a High-Resolution Green Fluorescent Protein Structure. *J. Am. Chem. Soc.* **2010**, *132* (32), 11093–11102.
- (29) Shinobu, A.; Agmon, N. Proton Wire Dynamics in the Green Fluorescent Protein. *J. Chem. Theory Comput.* **2017**, *13* (1), 353–369.
- (30) Chattoraj, M.; King, B. A.; Bubltz, G. U.; Boxer, S. G. Ultra-Fast Excited State Dynamics in Green Fluorescent Protein: Multiple States and Proton Transfer. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93* (16), 8362–8367.
- (31) Mardirossian, N.; Head-Gordon, M. Thirty Years of Density Functional Theory in Computational Chemistry: An Overview and Extensive Assessment of 200 Density Functionals. *Mol. Phys.* **2017**, *115* (19), 2315–2372.
- (32) Laptanok, S. P.; Lukacs, A.; Gil, A.; Brust, R.; Sazanovich, I. V.; Greetham, G. M.; Tonge, P. J.; Meech, S. R. Complete Proton Transfer Cycle in GFP and Its T203V and S205V Mutants. *Angew. Chem., Int. Ed.* **2015**, *54*, 9303–9307.
- (33) Voth, G. A. Computer Simulation of Proton Solvation and Transport in Aqueous and Biomolecular Systems. *Acc. Chem. Res.* **2006**, *39*, 143–150.

(34) Kaila, V. R. I.; Hummer, G. Energetics and Dynamics of Proton Transfer Reactions along Short Water Wires. *Phys. Chem. Chem. Phys.* **2011**, *13* (29), 13207–13215.

(35) Song, H.; van der Velden, N. S.; Shiran, S.; Bleiziffer, P.; Zach, C.; Sieber, R.; Imani, A. S.; Krausbeck, F.; Aebi, M.; Freeman, M. F.; et al. A Molecular Mechanism for the Enzymatic Methylation of Nitrogen Atoms within Peptide Bonds. *Sci. Adv.* **2018**, *4* (1–12), No. eaat2710.