A Theoretical Study on the Nature of On- and Off-States of Reversibly Photoswitching Fluorescent Protein Dronpa: Absorption, Emission, Protonation, and Raman

Xin Li,[†] Lung Wa Chung,[†] Hideaki Mizuno,[‡] Atsushi Miyawaki,[‡] and Keiji Morokuma*[,†]

Fukui Institute for Fundamental Chemistry, Kyoto University, Kyoto 606-8103, Japan, and Laboratory for Cell Function and Dynamics, Advanced Technology Development Group, Brain Science Institute, RIKEN, 2-1 Hirosawa, Wako-city, Saitama, 351-0198, Japan

Received: October 16, 2009

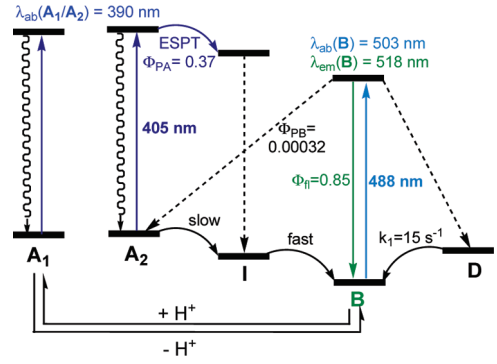
A GFP-like fluorescent protein, Dronpa, which was engineered from a coral Pectiniidae, was found to display perfect photochromic properties; the fluorescent "on"-state and nonfluorescent "off"-state of Dronpa can be reversibly switched by irradiation of two different wavelengths of light. To understand the detailed mechanism of the reversible photoswitching process at the atomic level, we performed QM and ONIOM(QM:MM) calculations to study the nature of the proposed on-state and off-state. Several high-level QM methods (TD-B3LYP, CASSCF, CASPT2, and SAC-CI) were employed to compute the vertical absorption and emission energies in the gas phase for four different protonation states as well as two conformations. The vertical absorption and emission energies of the on- and off-states in the proteins were further studied by the ONIOM(QM:MM) calculations. The ONIOM calculations on the absorption and emission suggest the neutral *trans* form is the off-state and the anionic *cis* form is the on-state. The dominant protonation states of the on- and off-states are also supported by protonation probability calculations via Poisson−Boltzmann electrostatics and Monte Carlo sampling. Moreover, the local protein environments were found to influence protonation states of the chromophore. Different possible reaction mechanisms are also discussed.

I. Introduction

Wild-type green fluorescent protein (GFP), which was discovered from *Aequorea victoria* jellyfish, and its variants play an indispensable role in biological imaging and analysis,[1,2] as recently honored by Nobel Prize in chemistry to three pioneers in the field. Photoactivatable fluorescent proteins, particularly reversibly photoswitching fluorescent proteins (RPFPs), have become a new class of fluorescent proteins (FPs).[3−7] The fluorescent on-state and nonfluorescent off-state of RPFPs can be reversibly switched by irradiation of two different radiations. Such a photochromic property of RPFPs advances the fluorescent protein technology, and can also be potentially applied for molecular switch and optical data storage.[8] Miyawaki et al. discovered one of the most promising RPFPs, Dronpa,[3a] which was engineered from a coral Pectiniidae. Protein dynamics *in vivo* (nucleocytoplasmic shuttling of signaling proteins) was successfully tracked by Dronpa.[3a] Miyawaki and co-workers also developed two mutants of Dronpa (Dronpa-2 and Dronpa-3) endowed with faster response to light and faster thermal relaxation from the off-state to the on-state.[9,10]

The absorption spectra of Dronpa show one major peak at 503 nm (2.46 eV) and one minor peak at 390 nm (3.18 eV) at pH 7.4.[3a] Decreasing pH diminishes the intensity of the former peak and increases that of the latter peak. The anionic (B) and neutral (A_1) forms of the chromophore have been assigned for the absorption peaks at 503 and 390 nm, respectively. The proposed anionic form is highly fluorescent ($\Phi_{fl} = 0.85$, $\lambda_{em} =$ 518 nm (2.39 eV), and $\tau_{fl} = 3.6$ ns). Interestingly, the photobleaching from the fluorescent state of Dronpa to the nonfluorescent state (A_2) was found to occur by intense excitation at 488 nm, with a quantum yield (Φ_{PB}) of 0.00032. The excited state of the A_2 form was found to undergo fast radiationless decay (with weak fluorescence of $\Phi_{fl} = 0.02$ and $\lambda_{em} = 450$ nm (2.76 eV), not in Scheme 1) but can be "activated" to the fluorescent state (B) by weak irradiation at 405 nm (photoactivation; $\Phi_{PA} = 0.37$).[3a] The reversible photoswitching could be repeated more than 100 times at the ensemble and single-molecule levels.[3] Although the absorption spectrum of the photoinduced nonfluorescent state A_2 is essentially identical to that of the proposed neutral form A_1 derived from acidification, the latter cannot be photoactivated to the fluorescent state B.[3a,b] On the basis of photophysical properties of Dronpa determined by the single-molecule spectroscopy, the reaction mechanism of the reversible photoswitching of Dronpa was proposed as shown in Scheme 1.[3] The excited-state proton transfer (ESPT) was suggested to proceed from the neutral A_2 form to give a nonfluorescent intermediate I, assumedly the

SCHEME 1: The Proposed Reaction Mechanism for the Reversible Photoswitching in Dronpa



* To whom correspondence should be addressed. E-mail: morokuma@fukui.kyoto-u.ac.jp.
[†] Kyoto University.
[‡] RIKEN.

10.1021/jp909947c
Published on Web 11/10/2009

deprotonated form in an unrelaxed protein environment, and eventually give to the anionic B form. This reaction mechanism is analogous to the three-state photoisomerization model for wild-type GFP.[1b,11] The ESPT in Dronpa was supported by the kinetic deuterium isotope effect (KIE $\sim$ 2).[3e] The unknown nonfluorescent and metastable D form was also proposed to account for the dynamic behavior of Dronpa.[3]

Recently, X-ray crystal structures of the on- and off-states of Dronpa were independently obtained by different groups.[12,13] The chromophore of Dronpa is formed by posttranslational modification from the Cys62-Tyr63-Gly64 (CYG) tripeptides. The chromophore is in a *cis* and coplanar conformation in the on-state crystal structures, while it was suggested to adopt a *trans* and nonplanar conformation in the off-state crystal structure.[12,13] In this connection, the chromophore for the nonfluorescent or dark state was also found to adopt a *trans* and nonplanar form in some GFP-like proteins.[2b,6a,b,14] The local environment around the chromophore was suggested to influence protonation states of the chromophore.[12d] As a result, the reaction mechanism involving *cis*−*trans* isomerization of the chromophore was proposed to dictate the protonation state and, in turn, the on/off-states, rather than the reaction mechanism initiated with ESPT (Scheme 1).

However, the detailed reaction mechanism of the reversible photoswitching in Dronpa at the atomic level remains unsolved. Characterizing the nature of the experimentally observed on- and off-states is of great importance in understanding the reaction mechanism of the photoswitching process and designing better molecular photoswitches. The assignment of the correct protonation state of the chromophore is challenging, and different protonation states have been proposed to be responsible for GFP. Although there are several theoretical studies on the wild-type GFP and other fluorescent proteins (e.g., asFP595),[15−17] no excited-state theoretical work on the most promising RPFP Dronpa has been reported.[18]

In this paper, systematic high-level ab initio QM and ONIOM(QM:MM) calculations were performed to delineate the nature of the on- and off-states of the chromophore by studying (1) the vertical absorption and emission energies of the chromophore with different possible protonation states in the gas phase (section III.A) and in the proteins (section III.B) and (2) the protonation probability of the chromophore in the protein (section III.C). Our calculations support the neutral *trans* form as the off-state and the anionic *cis* form as the on-state. The local environment, particularly Ser142, plays an important role in stabilizing the anionic *cis* form of the chromophore. The possibilities of different reaction mechanisms are also discussed (section III.D).

II. Computational Methods

A. System Preparation and Classical Simulations. The initial structures of the on- and off-states were obtained from the Protein Data Bank (PDB IDs: 2IOV and 2POX).[12b,d] Since Dronpa is monomeric in solution,[3a] chain A containing residues 1−218 of the proteins and surrounding crystal water molecules were used. The orientations of histidine, asparagine, and glutamine residues were examined by WhatCheck, MolProbity, and visual inspection to have better hydrogen bonding or less steric repulsion.[19] Missing hydrogen atoms were added and major hydrogen-bond networks were optimized by the program PDB2PQR.[20a] Meanwhile, protonation states of the titratable residues at pH 7 were determined by PROPKA implemented in PDB2PQR.[20] In addition, the nature of histidine, i.e., protonated at δ and/or ε nitrogen, was assigned on the basis of the local hydrogen-bonding network via visual inspection. Then, the prepared structures of the two proteins were fully solvated in the truncated octahedron water box constructed from a cubic box of 69.2 and 71.3 Å for 2IOV and 2POX, respectively, and neutralized by addition of one Cl^- counterion via the Amber Leap module.[21]

All classical molecular mechanics (MM) calculations have been performed with the AMBER all-atom force field and TIP3P water model.[22] All available force field parameters, charges, and atom types of each amino acid have been taken from the AMBER library. Since the atomic charges for the AMBER force field are derived from RESP charges calculated at the HF/6-31G(d) level, RESP atomic charges of the *cis* and *trans* forms of the chromophore were obtained at the same level by the Gaussian and Amber programs.[21,23] Additionally, similar Amber atom types derived from the Amber Antechamber module were applied for the chromophore. The water molecules and counterions were first optimized, and then, the entire system was further optimized to remove close contacts, followed by a series of molecular dynamics simulations. The chromophore in the two protein structures was kept frozen during the MM optimization and simulations. The simulations of 50 ps duration in the temperature ranges of 0−100, 100−200, and then 200−298 K under constant volume conditions have been performed to gradually heat the system. An additional 50 ps simulation in the NPT ensemble at 298 K and 1 bar was then performed. Subsequently, a 1.8 ns molecular dynamics simulation in the NVT ensemble was carried out. The simulations were conducted with periodic boundary conditions. A cutoff radius of 12 Å was used for nonbonding interactions, and electrostatic interactions were calculated with the particle mesh Ewald method.[24] The SHAKE algorithm was utilized to constrain all bond lengths involving hydrogen atoms.[25] The Leapfrog algorithm with a time step of 1 fs was used, and the temperature was controlled by the Langevin thermostat algorithm.[26] Finally, the last frame of the simulations, in which the solvation water molecules 8 Å or more away from the protein were deleted, was taken as the initial geometry of the ONIOM optimizations.

B. QM and ONIOM(QM:MM) Calculations. The gas-phase optimization for different protonation states of the truncated chromophore models, shown in Scheme 2, has been performed by B3LYP/6-31+G(d,p) and two-state-average SA2-CASSCF(14e,13o)/6-31G(d) (S_0:S_1 = 0.5:0.5) methods.[27,28] The vertical absorption and emission energies on the optimized structures were calculated by TD-B3LYP/6-31++G(d,p), CASPT2-(14e,13o)/6-31G(d), and SAC-CI(Level2)/D95(d) methods.[29,30] On the basis of our preliminary SA2-CASSCF(16e,14o)/3-21G calculations (with all π-electrons and orbitals in the active space), we adopted a large active space (14e,13o) for all calculations, which contains all π-electrons and orbitals except the amidic nitrogen lone-pair.[31] The molecular orbitals, occupancies, and major electronic configurations in the ground state and the first excited state of this large active space for the possible protonation states except for the neutral forms are shown in Figures S1−S6 in the Supporting Information. We also used a slightly smaller active space in the gas phase: (a) (12e,12o) for the protonated phenol forms [neutral (**N**) and cationic (**C**)] and (b) (12e,11o) for the deprotonated phenol forms [anionic (**A**) and zwitterionic (**Z**)].[15a] In general, as will be discussed later, the small active space leads to similar results for the large active space (Table 1 and Tables S1 and S2 in the Supporting Information). An average over three states (SA3; average of S_0, S_1, and S_2) was also performed in the gas phase, which gives similar results to those obtained by the two-state-

SCHEME 2: Possible Protonation States and Conformations

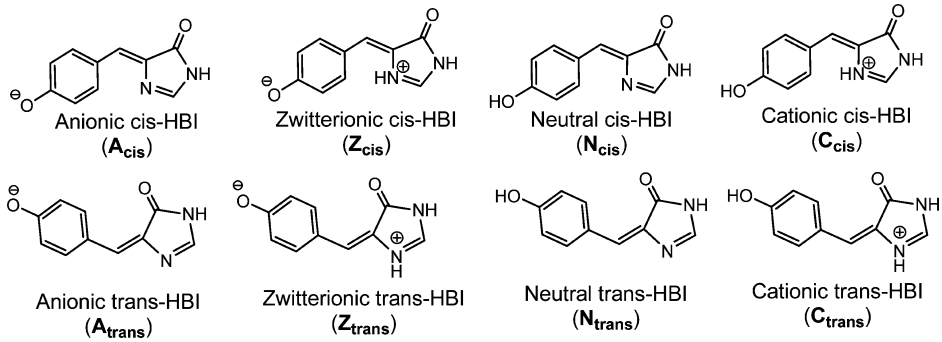


TABLE 1: Calculated Vertical Absorption Energies (in eV) and Oscillator Strengths (in Parentheses) of the Chromophore in Different Forms Calculated by the TD-B3LYP/6-31++G(d,p) and SAC-CI(Level2)/D95(d) Methods at the B3LYP/6-31+G (d,p)-Optimized Structures

conformation:	*cis*		*trans*	
method:	TD-B3LYP	SAC-CI	TD-B3LYP	SAC-CI
anionic, **A**	3.05(0.90)	2.48(0.89)	3.11(0.85)	2.48(0.79)
zwitterionic, **Z**	2.88(0.73)	2.16(0.64)	3.05(0.85)	2.68(0.89)
neutral, **N**	3.48(0.70)	3.49(0.76)	3.46(0.56)	3.38(0.59)
cationic, **C**	3.11(0.67)	2.64(0.65)	3.16(0.69)	2.59(0.66)
Expt.	gas, $\mathbf{A}_{cis}$ (2.59); toluene, $\mathbf{A}_{cis}$ (2.79) and $\mathbf{N}_{cis}$ (3.43)			

average (SA2) calculations (Table S1 in the Supporting Information). A level shift of 0.3 au was used in the CASPT2 calculations in order to avoid intruder state problems.[32] All CASSCF and CASPT2 calculations were performed by the Molpro program unless otherwise stated.[33] The Raman spectra of the essential protonation states have been calculated by the B3LYP/6-31+G(d,p) method with the suggested scaling factor.[34]

As recently reported,[35] we found that the situation for the neutral forms is rather complex as resulting from CASSCF calculations. The desired excitation or emission state was found to be of type S_3 in the four-state-average SA4-CASSCF (average of S_0, S_1, S_2, and S_3) level. S_3 has considerable dynamic correlation and becomes lower in energy than the other two electronic excited states in CASPT2 calculations. Thus, the SA2-CASSCF or SA3-CASSCF method cannot describe the desired excited state. Therefore, optimization of the S_3 state for the neutral form was performed by the SA4-CASSCF(12e,12o)/6-31G(d) and SA4-CASSCF(14e,13o)/6-31G(d) methods for the emission.[36] The calculated vertical absorption and emission energies were evaluated by singlet-state (SS) and multistate (MS) CASPT2(12e,12o) or CASPT2(14e,13o) (including S_0, S_1, S_2, and S_3) with ANO-S basis sets ([4s3p1d] for all elements except [2s1p] for the hydrogen) and an imaginary level shift of 0.3 au, using the Molcas7 program.[37] Two different zero-order Hamiltonians for the neutral forms were tested. Compared to the experimental value, the latest IPEA Hamiltonian[38] was found to have a larger error of the vertical absorption energy than the original Hamiltonian. Also, compared to SS-CASPT2, the effect of MS-CASPT2 was found to be quite large for some calculations by using the IPEA zero-order Hamiltonian. Therefore, the results obtained from the original Hamiltonian are mainly discussed and detailed results for both Hamiltonians are given in the Supporting Information.

The protein structures containing the *cis* and *trans* forms of the chromophore prepared from the classical simulations were refined by the ONIOM(QM:MM) optimization for different protonation states in order to compute the vertical absorption and emission energies.[39,40] In the ONIOM calculations, the QM model part is calculated by the above-mentioned QM methods and the MM part is treated by the AMBER force field.[22] All ONIOM(QM:MM) calculations were conducted by the Gaussian program.[23] For the case of the ONIOM optimization with CASSCF as the QM method, QM energies and gradients were computed externally by the Molpro program.[33]

In two-layer ONIOM(QM:MM) calculations, the interface between the QM and MM regions is treated by hydrogen link atoms. The total energy of the system in ONIOM(QM:MM) is calculated by the following equation:

$$E_{\mathrm{ONIOM}} = E_{\mathrm{MM,real}} + E_{\mathrm{QM,model}} - E_{\mathrm{MM,model}} \qquad (1)$$

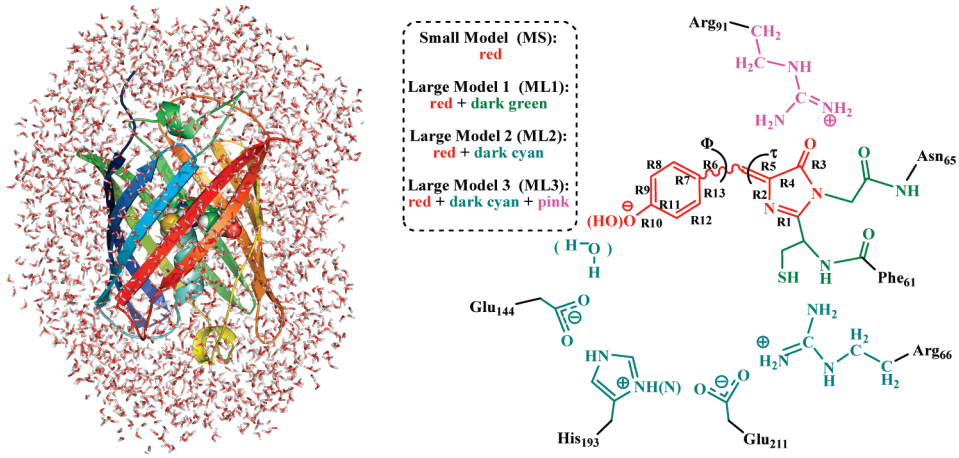
Here, $E_{\mathrm{MM,real}}$ is the MM energy of the entire system, called the real system in the ONIOM method; $E_{\mathrm{QM,model}}$ is the QM energy of a major chemically interested part of the real system, called the model part; and $E_{\mathrm{MM,model}}$ is the MM energy of the model part. The real system was further divided into the optimized MM region and the frozen MM region.[41] The residues within about 9 Å of the chromophore were assigned to the optimized MM region and allowed to be optimized.

When performing ONIOM calculations, one can treat the interaction of the MM point charges with the QM part in the MM (called mechanical embedding or ME scheme) or in the QM Hamiltonian (called electronic embedding or EE scheme). In the EE scheme, the wave function of the QM part is polarized by the MM point charges. Geometry optimization has been performed in both ME and EE schemes. For all absorption and emission calculations in the proteins, only the EE scheme has been used. To avoid overpolarization of the QM wave function, the Gaussian default scaling parameters for the boundaries have been used.[42] It is noted that SAC-CI and CASPT2 cannot be used as the QM method in the ONIOM(QM:MM) calculations in the present Gaussian implementation. However, one recognizes, for single point calculations of excitation or emission energies, the MM contributions cancel out and only the interaction of the MM point charges with the QM Hamiltonian needs to be evaluated: this can be accomplished by running SAC-CI and CASPT2 calculations for ground and excited states in the presence of the MM charges without using the ONIOM code. The results have been confirmed to be true for ONIOM(TD-B3LYP:AMBER)-EE calculations. The excitation and emission energies we obtained using this method will be reported later as ONIOM(SAC-CI:AMBER)-EE or ONIOM(CASPT2:AMBER)-EE results.

Several QM models shown in Scheme 3[43] have been adopted in our ONIOM calculations, in order to examine the effects of including some protein residues in QM calculations.

C. Protonation Probability Calculations. The protonation probability of the on- and off-state structures was evaluated with

SCHEME 3: Structure and QM Models in the ONIOM Calculations



continuum electrostatic calculations by solving the linear Poisson−Boltzmann (LPB) equation via the program MEAD from Bashford and Karplus[44] and sampling the ensemble of protonation patterns by a Monte Carlo (MC) algorithm via the program Karlsberg.[45] The ESP charges are used for the chromophore at the different protonation states, while the Parse charges are used for the other residues in the protein.[46] In addition, the Parse radii have been used for the calculations.[46] A two-step grid-focusing procedure was used to solve the LPB equation with grid resolutions of 1.0 and 0.25 Å. All water molecules were removed in the electrostatic energy calculations.[47] The dielectric constants were set to 4 for the protein and 80 for the water solvent. The solute−solvent boundary was defined through a solvent probe radius of 1.4 Å. The ionic strength is 0.12 M, and the temperature is 300 K. Standard pK_a values of model compounds were used for the side-chains of the studied residues: Glu (4.4) and His (6.6).[48] The pK_a value of the CYG imidazolinone nitrogen has been evaluated in accordance with the thermodynamic cycle 1 described in the literature[49] by the B3LYP/6-31+G(d,p) method using the conductor-like polarizable continuum model[50] in water (3.1 for $\mathbf{Z}_{cis}$),[51] while the pK_a value of 8.2 measured for the GFP chromophore analogue[52] was used for that of the phenol moiety. The MC calculations involved 15 000 full and 25 000 reduced scans. Double (triple) MC steps have been applied to groups having an interaction energy larger than 2 (3) pK units. The pH ranges from −10.0 to 20.0, with an increase of pH by 0.1 units during the MC simulations.

III. Results and Discussion

A. Absorption and Emission in the Gas Phase. The calculated geometries and vertical absorption and emission energies for the truncated chromophore in the gas phase were summarized in Figure 1 and Tables 1−2, as well as in Figures S9−S16 and Tables S1−S3 of the Supporting Information.

A.1. Geometrical Features. Before we discuss the calculated absorption and emission in detail, we comment on the main geometrical features of the chromophore. The calculated key bond lengths of the chromophores are generally not much different between the *cis* and *trans* conformations both for the ground state and excited state, as shown in Figure 1 (also Figure S15 in the Supporting Information). Also, the ground-state geometries calculated by the B3LYP/6-31+G(d,p) method are quite similar to those by the SA2-CASSCF(14e,13o)/6-31G(d) method (Figures S11 and S12 in the Supporting Information), with a maximum deviation of about 0.03 Å. The B3LYP-optimized ground-state geometry should be more reliable, due to a lack of dynamic correlation in the CASSCF method.[40e,53] The geometrical deviation presumably results in somewhat different absorption energies (Tables 1 and 2), which will be discussed later. Owing to delocalization of the π-electrons, the ground-state anionic and zwitterionic forms have considerable quinone character. Therefore, the anionic forms have similar bond lengths for the two bridging bonds R5 and R6 (defined in Scheme 3) and, particularly, the zwitterionic forms have a longer bond length for R5 than R6, whereas the ground-state neutral and cationic forms have significant C═C character for R5 and

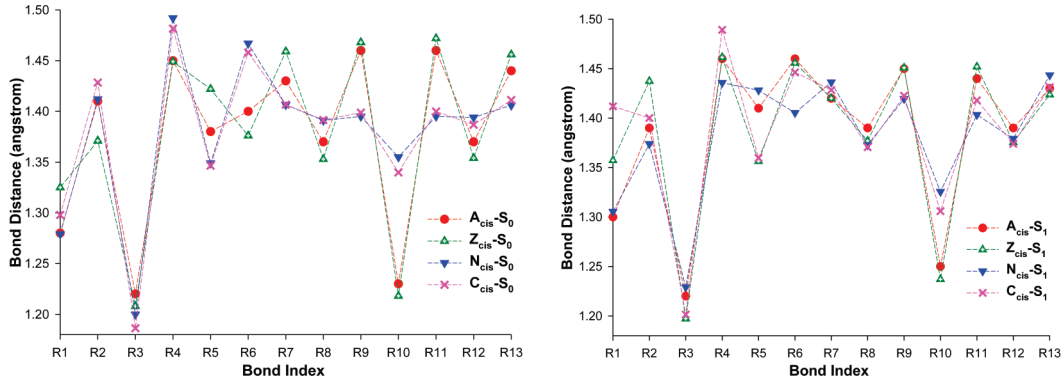


Figure 1. Key bond lengths of the ground-state (left) and excited-state (right) *cis* chromophore at the different protonation states in the gas phase calculated by the SA2-CASSCF(14e,13o)/6-31G(d) method, except for the excited-state $\mathbf{N}_{cis}$ by the SA4-CASSCF(12e,12o)/6-31G(d) method. The labeling is given in Scheme 3.

TABLE 2: Vertical Absorption and Emission Energies (in eV) of the Chromophore in Different Forms Calculated by the CASPT2(14e,13o)/6-31G(d) and SAC-CI(Level2)/D95(d) Methods at the SA2-CASSCF(14e,13o)/6-31G(d)-Optimized Structures, Except for the Neutral Forms

conformation:	*cis*		*trans*	
method:	CASPT2[a]	SAC-CI	CASPT2[a]	SAC-CI
		Absorption		
anionic, **A**	2.70	2.55	2.71	2.56
zwitterionic, **Z**	3.07	2.44	3.41	2.72
neutral, **N**	3.62[b,c]	3.67	3.69[b,d]	3.63
cationic, **C**	4.03	3.00	2.81	2.75
		Emission		
anionic, **A**	2.46	2.29	2.48	2.29
zwitterionic, **Z**	1.53	0.86	2.10	1.20
neutral, **N**	2.98[e]	2.91[e]	2.78[e,f]	2.74[e,g]
cationic, **C**	2.24	1.39	2.54	2.33

[a] The original zero-order Hamiltonian was used for all SS- and MS-CASPT2 calculations, and MS-CASPT2 results obtained from the IPEA Hamiltonian for the neutral forms were given in Table S3 of the Supporting Information. [b] It is obtained by the MS-CASPT2(12e,12o)/ANO-S calculation based on the SA2-CASSCF(12e,12o)/6-31G(d)-optimized geometry. [c] The calculated excitation energy is 3.49 eV by the MS-CASPT2(14e,13o)/ANO-S calculation based on the SA2-CASSCF-(12e,12o)/6-31G(d)-optimized geometry. [d] The calculated excitation energy is 3.31 eV by the MS-CASPT2(14e,13o)/ANO-S calculation based on the SA2-CASSCF(12e,12o)/6-31G(d)-optimized geometry. [e] It is obtained by the MS-CASPT2(12e,12o)/ANO-S calculation based on the SA4-CASSCF(12e,12o)/6-31G(d)-optimized geometry. [f] The calculated emission energy is 2.54 eV by the MS-CASPT2(14e,13o)/ANO-S calculation based on the SA4-CASSCF(14e,13o)/6-31G(d)-optimized geometry. [g] The calculated emission energy is 2.76 eV by the SAC-CI(Level2)/D95(d) calculation based on the SA4-CASSCF-(14e,13o)/6-31G(d)-optimized geometry.

C—C character for R6. These geometrical features can be attributed to different resonance structures (Scheme S1 in the Supporting Information).

Geometry optimizations with the CASSCF method can compare geometrical changes of the chromophores from the ground state to the excited state. Due to the intramolecular charge transfer excitation, some of the key bonds, especially R5 and R6, are altered in the excited state (Figure S16 in the Supporting Information). Such bond alternation for R5 and R6 should affect energetic profiles for photoisomerization along these two bonds. (a) In the excited-state anionic forms ($\mathbf{A}_{cis}$ and $\mathbf{A}_{trans}$), both R5 and R6 are elongated by 0.03 and 0.06 Å, respectively. (b) Considerable bond alternation is found in the excited-state $\mathbf{Z}_{cis}$ (R2, R5, and R6 by more than 0.06 Å) and $\mathbf{Z}_{trans}$ (R1 and R6 by more than 0.07 Å).[54] As a result, R6 becomes longer than R5. (c) Excitation of the neutral forms ($\mathbf{N}_{cis}$ and $\mathbf{N}_{trans}$) elongates R5 by about 0.08 Å, whereas R6 is shortened by 0.06 Å and, thus, becomes similar in bond length to R5 in the excited state. Several resonance structures are suggested to be responsible for the above-mentioned geometrical features (Scheme S2 in the Supporting Information).

A.2. Absorption. In the experiments, the absorption energies of the *cis* synthetic GFP analogue chromophores for the anionic, cationic, and neutral form are 2.79, 3.10, and 3.43 eV in the nonpolar toluene solution, respectively.[51,55] In the gas phase, the measured absorption energies for the *cis* anionic and cationic GFP analogues are 2.59 and 3.05 eV, respectively.[56]

The B3LYP Geometries. First, the vertical absorption energies of the chromophores evaluated by the TD-B3LYP/6-31++G(d,p) method and the more accurate SAC-CI(Level2)/D95(d)[30b−d] method at the B3LYP/6-31+G(d,p)-optimized geometries are discussed (Table 1). The absorption energy for the *cis* anionic ($\mathbf{A}_{cis}$: 2.48 eV) and neutral ($\mathbf{N}_{cis}$: 3.49 eV) forms obtained by the SAC-CI method is quite close to that for the analogue chromophores measured in the gas phase or toluene.[51,56] However, the TD-B3LYP method, which fails to give a good description of the charge transfer state,[57] is found to have an error of about 0.46 eV for the anionic state,[15l−n,17a,56b] while it can reproduce the absorption energy for the neutral form.[58] Both the TD-B3LYP and SAC-CI calculations suggest that most of the *trans* chromophores have similar absorption energies to the corresponding *cis* chromophores except for $\mathbf{Z}_{cis}$. Both methods give a much lower absorption energy for $\mathbf{Z}_{cis}$ than for $\mathbf{Z}_{trans}$. We trust the higher-level SAC-CI method, which has been shown to be very reliable in calculations of excitation energies of many molecular systems[30] and in the present system successfully reproduces the excitation energy of the anionic and neutral forms, and this method will thus be focused on in all subsequent calculations.[59] The theoretically lower-level TD-B3LYP method, which gives a similar absorption energy for all protonation states and cannot definitely assign the protonation state for the on- and off-state proteins, particularly for the anionic form, will not be used for further discussions (the TD-B3LYP results can be found in the Supporting Information).

The CASSCF Geometries. We also evaluated the absorption energy by the CASPT2 and SAC-CI methods at the CASSCF-optimized geometries (Table 2 and Table S2 in the Supporting Information). There are some minor differences (e.g., in $\mathbf{C}_{cis}$) in the SAC-CI absorption energies between Tables 1 and 2, due to a small difference in optimized geometries between B3LYP and CASSCF optimization (Figure S11 in the Supporting Information). CASPT2 absorption energies, as well as SAC-CI energies, are similar between *cis* and *trans* conformations.[60] Comparing the computed absorption energies with the experiment, the MS-CASPT2 calculated $\mathbf{N}_{cis}$ absorption energy and the CASPT2 calculated $\mathbf{A}_{cis}$ absorption energy agree with the experimental values.[51,56b,61] In addition, CASPT2 suggests that $\mathbf{Z}_{cis}$ has a higher absorption energy than $\mathbf{A}_{cis}$, but SAC-CI gives similar energies between these two forms. A higher absorption energy for the zwitterionic form than the anionic form has recently been reported by aug-MCQDPT2/CASSCF(16e,14o)//PBE0/(aug)-cc-pVDZ calculations, which includes dynamic correlation, on the asFP595 chromophore.[15m]

A.3. Emission. In order to further characterize the nature of the on- and off-states of Dronpa, we also calculated the vertical emission energy of the chromophore in different forms in the gas phase by the SAC-CI and CASPT2 methods at the CASSCF-optimized excited-state geometries. Both the CASPT2 and SAC-CI calculations give the emission energy for $\mathbf{A}_{cis}$ (2.46 and 2.29 eV, respectively) and for $\mathbf{N}_{trans}$ (2.78 and 2.74 eV, respectively) in good agreement with the experimental values of 2.39 and 2.76 eV for the on- and off-states of Dronpa.[3a] The CASPT2 and SAC-CI calculations also show that the *cis* and *trans* conformers have similar emission energies in all of the forms, although differences are more noticeable in zwitterionic and cationic forms.[54] It should be noted that $\mathbf{Z}_{cis}$ shows a significant change in geometry upon excitation and has a much lower emission energy than the experiment for the on-state of Dronpa.

A.4. The Nature of Excitation and Emission. Analysis of CASSCF (Figures S1−S10 in the Supporting Information) as well as TD-B3LYP and SAC-CI wave functions indicates clearly that the lowest excitation and emission for all forms of chormophore mainly involve $\pi \rightarrow \pi^*$ transition and $\pi^* \rightarrow \pi$ transition, respectively. Since the ground-state and excited-state

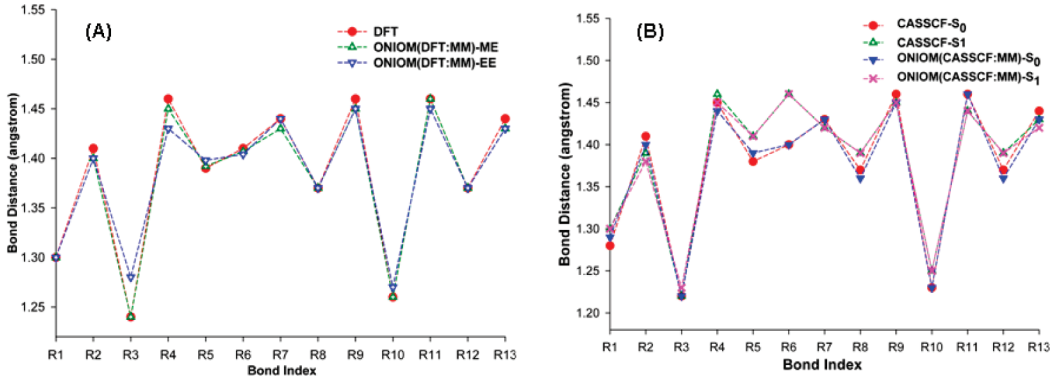


Figure 2. Calculated key bond lengths of $\mathbf{A}_{cis}$: (A) the ground-state chromophore in the gas phase by the B3LYP/6-31+G(d,p) method and in the protein by the ONIOM(B3LYP/ 6-31+G(d,p):AMBER)-ME and -EE methods; (B) the ground-state (S_0) and first excited-state (S_1) chromophore in the gas phase by the SA2-CASSCF(14e,13o)/6-31G(d) method and in the protein by the ONIOM(SA2-CASSCF(14e,13o)/6-31G(d):AMBER)-ME method.

geometries give very similar wave functions, only analyses of the excitation based on the CASSCF method are summarized as follows. (a) The lowest excitation of the anionic forms, $\mathbf{A}_{cis}$ and $\mathbf{A}_{trans}$, involves intramolecular charge transfer mainly from the phenoxide and imidazolinone moieties to the bridge CH moiety (~0.07−0.09 e). (b) The zwitterionic forms, $\mathbf{Z}_{cis}$ and $\mathbf{Z}_{trans}$, are found to have considerable multiconfigurational character in both the ground and excited states. Therefore, the single-configurational-based methods, e.g., the TD-DFT method, could not give the correct description for the zwitterionic chromophores. In $\mathbf{Z}_{cis}$, about 0.05 e is transferred upon excitation from the bridge part mainly to the phenoxide moiety, while, in $\mathbf{Z}_{trans}$, approximately 0.07 e is transferred upon excitation from the bridge and phenoxide parts to the imidazolinone moiety and a small contribution of a double excitation configuration can be seen. (c) Different from the considerable delocalization in the anionic (**A**) and zwitterionic (**Z**) chromophores, the lowest excitation of the neutral chromophores, $\mathbf{N}_{cis}$ and $\mathbf{N}_{trans}$, involves a large amount of charge transfer mainly from the phenol ring to the imidazolinone part (~0.3 e). (d) Significant intramolecular charge transfer from the phenol ring to the bridge and imidazolinone parts is also observed (~0.11−0.45 e) in the lowest excitation of $\mathbf{C}_{cis}$ and $\mathbf{C}_{trans}$.

In summary, the vertical absorption and emission energies as well as the nature of excitation in different conformations and protonation states have been discussed. On the basis of the SAC-CI absorption energy, $\mathbf{A}_{cis}$, $\mathbf{Z}_{cis}$, and $\mathbf{C}_{cis}$ are found to give an absorption energy similar to the experiment for the on-state of Dronpa. However, according to the experimental fact, the intensity of the absorption at 503 nm is reduced by decreasing pH,[3a] which should eliminate the possibility of $\mathbf{C}_{cis}$ to be the on-state of Dronpa. Additionally, the calculated emission energies for $\mathbf{Z}_{cis}$ and $\mathbf{C}_{cis}$ are not consistent with the experiment. Thus, SAC-CI results clearly indicate that $\mathbf{A}_{cis}$ is the on-state of Dronpa. This conclusion is also supported by the CASPT2 calculations. On the other hand, the SAC-CI and CASPT2 calculated absorption energies in the gas phase for $\mathbf{N}_{trans}$ and for $\mathbf{C}_{trans}$ are higher and lower, respectively, than the experimental value (absorption: 3.18 eV) for the off-state of Dronpa.[3a] Such a deviation between the gas-phase calculations and the protein experiment could be resolved by including the protein effect (section III.B). Although the computed absorption and emission energies in the gas phase provide a qualitative trend, definite assignments of the protonation state for the on-state and, particularly, the off-state of Dronpa are still not conclusive at this stage. It is the subject of the following ONIOM studies by including the proteins.

B. Absorption and Emission in the Proteins. The effects of protein environment on the absorption and emission energies for the on- and off-states of Dronpa have been examined by the ONIOM(QM:MM) calculations. In the following three sections, the geometries, the absorption and emission energies, and the protein effects for the on- and off-state Dronpa will be discussed.

B.1. Geometrical Features. We optimized the two protein structures (the on-state protein with the *cis* chromophore and the off-state protein with the *trans* chromophore) in four different protonation states, using the ONIOM(B3LYP:AMBER) approach with both mechanical embedding (ME) and electronic embedding (EE) schemes.

The main structural features of the possible chromophores in the on-state and off-state Dronpa (i.e., $\mathbf{A}_{cis}$ and $\mathbf{N}_{trans}$) are first summarized. The structures of these ground-state chromophores in the respective proteins optimized with ONIOM(B3LYP:AMBER) are compared with those optimized with B3LYP in the gas-phase (Figure 2A and Figure S17 in the Supporting Information). The ONIOM-ME structure is in general very similar to the gas-phase structure. The major difference is that, while the chromophore for both $\mathbf{A}_{cis}$ and $\mathbf{N}_{trans}$ is coplanar in the gas phase, the chromophore for $\mathbf{N}_{trans}$ in the off-state protein becomes substantially nonplanar with the dihedral angles around the two exocyclic bonds, ϕ and τ, of the chromophore (defined in Scheme 3) to be 13.9 and −0.4°, respectively. The nonplanarity of the chromophore mainly results from ϕ, the rotation around the C−C bond (R6). The chromophore for $\mathbf{A}_{cis}$ in the on-state protein is more coplanar ($\phi = -7.6°$ and $\tau = -178.1°$). Compared to the ONIOM-ME optimizations, the ONIOM-EE optimizations[62] or using a larger QM model for the chromophore[63] (ML1 in Scheme 3) leads to minor geometrical changes of the chromophore only. Therefore, the small QM model MS with the ONIOM-ME scheme was shown to describe the main geometrical features of the on- and off-state Dronpa in the ground state, and will be used for most of our ONIOM optimizations.

The excited-state chromophores in the on- and off-state proteins were further optimized by the ONIOM(CASSCF:AMBER)-ME method. As shown in Figure 2B (also Figure S17 in the Supporting Information), the calculated key bond lengths of the ground- and excited-state chromophores in the proteins are generally similar to the gas phase. The chromophores in the ground state and excited state are coplanar in the gas phase.[54] However, the ONIOM(CASSCF:AMBER)-ME-optimized ground-state chromophore for $\mathbf{N}_{trans}$ is highly noncoplanar ($\phi = 29.9°$ and $\tau = -3.6°$) in the off-state protein, and the

TABLE 3: Calculated Vertical Absorption Energies (in eV) and Oscillator Strengths (in Parentheses) of the Chromophore in the *cis* Conformation (On-State Protein) and in the *trans* Conformation (Off-State Protein) in Different Protonation States Calculated by the ONIOM(SAC-CI(Level2)/D95(d):AMBER)-EE Method at the ONIOM(B3LYP/6-31+G(d,p):AMBER)-EE-Optimized Ground-State Structures

conformation in proteins:	*cis* (on-state)		*trans* (off-state)	
QM model:	MS	ML1	MS	ML1
anionic, **A**	2.36(0.85)	2.26(0.83)	2.24(0.49)	2.21(0.71)
zwitterionic, **Z**	2.21(0.40)	2.11(0.70)	2.73(0.69)	2.35(0.84)
neutral, **N**	3.09(0.73)	3.27(0.64)	3.01(0.56)	3.03(0.51)
cationic, **C**	2.42(0.71)	2.80(0.70)	2.36(0.58)	2.57(0.61)
Expt.	2.46		3.18	

nonplanarity is decreased to 19.9° (ϕ) and −2.5° (τ) in the excited state. The calculated values for nonplanarity are smaller than the X-ray crystal structure of $\mathbf{N_{trans}}$ ($\phi_{aver} = 40.1°$ and $\tau_{aver} = -20.7°$), possibly due to crystal packing and slightly different rotation barriers in different methods. The ground-state chromophore for $\mathbf{A_{cis}}$ is nearly planar ($\phi = -6.2°$ and $\tau = 179.5°$) in the on-state protein, whereas it becomes more noncoplanar in the excited state ($\phi = -20.5°$ and $\tau = -174.8°$). The nonplanarity of the chromophore is presumably attributed to the bond strength of R5 and R6 (affecting τ and ϕ, respectively, see Figure 2 and Figure S17 in the Supporting Information), and also the immediate protein environment.

B.2. Absorption and Emission. As mentioned in the Introduction, the recent X-ray crystal structures suggest that the chromophore adopts the *cis* conformation in the on-state Dronpa and the *trans* conformation in the off-state Dronpa.[12,13] The absorption energies of the chromophore in the four protonation states in the two protein structures were calculated, as shown in Table 3, by the ONIOM(SAC-CI:AMBER)-EE method[64] at the ONIOM(B3LYP:AMBER)-EE-optimized structures.[65,66]

In general, the ONIOM SAC-CI calculations show that the small and large QM models (MS and ML1) for the chromophore have similar absorption energies for the same protonation states, regardless of the on- or off-state protein (Table 3). The calculated absorption energy for $\mathbf{A_{cis}}$ (2.26−2.36 eV) is in good agreement with the experimental value for the on-state Dronpa (2.46 eV).[3a] Nonetheless, the absorption energy for $\mathbf{Z_{cis}}$ is also close to the experimental one.[3a] The possibility of $\mathbf{Z_{cis}}$ being the on-state of Dronpa will be further examined by the protonation probability calculations (section III.C). The calculated absorption energy for the nonplanar $\mathbf{N_{trans}}$ in the off-state protein (3.01−3.03 eV) is consistent with the experimental one (3.18 eV), while those for other protonated states are far from the experiment.[3a] In addition, $\mathbf{N_{trans}}$ is further supported by the ONIOM TD-B3LYP calculations (3.33 eV for MS and 3.17 eV for ML1, see Table S4 in the Supporting Information), in which the TD-B3LYP can well describe the neutral form (Table 1).[66]

The detailed protein effects on the absorption energy of Dronpa will be further investigated. At first, the effects of including nearby key charged residues in the QM part on the absorption of $\mathbf{A_{cis}}$ and $\mathbf{N_{trans}}$ were studied by enlarging the QM models (ML2 and ML3, see Scheme 3). The ONIOM SAC-CI calculated absorptions are insignificantly influenced by this QM effect of the nearby residues ($\mathbf{A_{cis}}$, 2.35 eV for ML2 and 2.45 eV for ML3; $\mathbf{N_{trans}}$, 3.07 eV for ML2 and 3.03 eV for ML3), which are similar to those for MS and ML1. Therefore, the ONIOM SAC-CI absorption energy calculations with small QM model MS at the ONIOM-ME-optimized protein structure[65] are generally sufficient to study the absorption and will be used in the following calculations.

TABLE 4: Absorption and Emission Energies (in eV) of $\mathbf{A_{cis}}$, $\mathbf{Z_{cis}}$, and $\mathbf{N_{trans}}$ in the Proteins Calculated by the ONIOM(CASPT2(14e,13o)/6-31G(d):AMBER)-EE and ONIOM(SAC-CI (Level2)/D95(d):AMBER)-EE Methods at the ONIOM(CASSCF(14e,13o)/6-31G(d):AMBER)-ME-Optimized Ground-State and Excited-State Structures, except for $\mathbf{N_{trans}}$

method:	CASPT2[a]	SAC-CI	Expt.
	Absorption		
anionic *cis*, $\mathbf{A_{cis}}$	2.71	2.42	2.46
zwitterionic *cis*, $\mathbf{Z_{cis}}$	3.29	2.54	
neutral *trans*, $\mathbf{N_{trans}}$	3.24[b,c]	3.18	3.18
	Emission		
anionic *cis*, $\mathbf{A_{cis}}$	2.42	2.08	2.39
zwitterionic *cis*, $\mathbf{Z_{cis}}$	1.63	0.67	
neutral *trans*, $\mathbf{N_{trans}}$	2.84[d,e]	2.12[d]	2.76

[a] The original zero-order Hamiltonian was used for all ONIOM SS- and MS-CASPT2 calculations, and ONIOM MS-CASPT2 results obtained from the IPEA Hamiltonian were given in Table S5 of the Supporting Information. [b] It is obtained from the ONIOM MS-CASPT2(12e,12o)/ANO-S method based on the ONIOM SA2-CASSCF(12e,12o)/6-31G(d)-optimized geometry. [c] The calculated excitation energy is 3.16 eV by the ONIOM MS-CASPT2(14e,13o)/ANO-S method based on the ONIOM SA2-CASSCF(12e,12o)/6-31G(d)-optimized geometry. [d] It is obtained from the ONIOM MS-CASPT2(12e,12o)/ANO-S method based on the ONIOM SA4-CASSCF-(12e,12o)/6-31G(d)-optimized geometry. [e] The calculated emission energy is 2.92 eV by the ONIOM MS-CASPT2(14e,13o)/ANO-S method based on the ONIOM SA4-CASSCF(12e,12o)/6-31G(d)-optimized geometry.

In order to gain more insight, the ONIOM(CASSCF:AMBER)-ME optimizations for the most likely forms ($\mathbf{A_{cis}}$, $\mathbf{Z_{cis}}$, and $\mathbf{N_{trans}}$) in the on- and off-states of Dronpa have been further performed, followed by the ONIOM(CASPT2:AMBER)-EE and ONIOM(SAC-CI:AMBER)-EE single-point calculations to evaluate emission as well as absorption energies (Table 4). Both the ONIOM CASPT2 and SAC-CI calculations support $\mathbf{A_{cis}}$ as the on-state of Dronpa, although the CASPT2 calculation slightly overestimates the absorption energy. The ONIOM CASPT2 calculations, which should take better care of the multiconfigurational nature of the excited state of $\mathbf{Z_{cis}}$, do not support $\mathbf{Z_{cis}}$ as the on-state. In addition, both the ONIOM MS-CASPT2[67] and SAC-CI methods (here at the ONIOM(CASSCF:AMBER)-ME-optimized structure) reproduce the experimental absorption for $\mathbf{N_{trans}}$ in the off-state protein.[3a]

Furthermore, the vertical emission energies in the proteins were also examined (Table 4), in which the excited-state chromophore in the proteins was optimized by the ONIOM-(CASSCF:AMBER)-ME method. The calculated emission energy for $\mathbf{A_{cis}}$ by the ONIOM SAC-CI method was found to be underestimated by 0.31 eV, while that by the CASPT2 method is in good agreement with the experimental value with a deviation of about 0.03 eV. Both the CASPT2 and SAC-CI methods lead to a much lower emission energy for $\mathbf{Z_{cis}}$, in which the excited-state chromophore is slightly deformed, particularly around the cationic nitrogen atom.[54] These results should further eliminate the possibility of $\mathbf{Z_{cis}}$ as the on-state chromophore of Dronpa. For the case of the off-state Dronpa, the ONIOM MS-CASPT2 calculated emission energy for $\mathbf{N_{trans}}$ agrees well with the experimental value, but ONIOM SAC-CI underestimates the emission, as shown in Table 4.

B.3. Origin of the Protein Effects. The effects of protein on the absorption and emission of the chromophore can be

TABLE 5: Protein Effects for the ONIOM(SAC-CI(Level2)/D95(d):AMBER)-EE and ONIOM(CASPT2(14e,13o)/6-31G(d):AMBER)-EE Absorption and Emission Energies (in eV) for A_{cis} and N_{trans} (QM Model MS) in the Proteins at the ONIOM(B3LYP/6-31+G(d,p):AMBER)-EE- and ONIOM(SA2-CASSCF(14e,13o)/6-31G(d):AMBER)-ME-Optimized Geometries, Except for N_{trans}

	QM for energy//QM for geometry	gas phase	protein: geometric	protein: total
		ΔE_{gas} [a]	$\Delta E_{QM,model}$ [b]	$\Delta E_{ONIOM\text{-}EE}$
absorption for $\mathbf{A_{cis}}$	SAC-CI//B3LYP	2.48	2.31	2.36
	SAC-CI//CASSCF	2.55	2.33	2.42
	CASPT2//CASSCF	2.70	2.69	2.71
emission for $\mathbf{A_{cis}}$	SAC-CI//CASSCF	2.29	1.96	2.08
	CASPT2//CASSCF	2.46	2.39	2.42
absorption for $\mathbf{N_{trans}}$	SAC-CI//B3LYP	3.38	3.09	3.01
	SAC-CI//CASSCF	3.63	3.39	3.18
emission for $\mathbf{N_{trans}}$	SAC-CI//CASSCF[c]	2.74	2.62	2.12

[a] The calculated energy difference between S_0 and S_1 in the gas phase. [b] The calculated energy difference between S_0 and S_1 in the gas phase for the QM model part taken from the ONIOM-optimized geometry. [c] The SA4-CASSCF(12e,12o)/6-31G(d)-optimized geometry is used.

TABLE 6: Calculated Population (%) of Different Protonation States at pH 7[a]

	off-state $\mathbf{A_{trans}}$ vs $\mathbf{N_{trans}}$		on-state $\mathbf{A_{cis}}$ vs $\mathbf{N_{cis}}$		on-state $\mathbf{A_{cis}}$ vs $\mathbf{Z_{cis}}$		
geometry	ONIOM//$\mathbf{N_{trans}}$[b]	X-ray	ONIOM//$\mathbf{A_{cis}}$[c]	X-ray	ONIOM//$\mathbf{A_{cis}}$[c]	X-ray	ONIOM//$\mathbf{Z_{cis}}$[d]
chromophore	N: 100	N: 100	A: 100	A: 94	A: 98	A: 100	A: 100
Glu144	A: 100	A: 100	A: 100	A: 100	A: 100	A: 100	A: 100
His193	HIE: 100	HIE: 100	HIP: 99	HIP: 49	HIP: 98	HIP: 43	HIP: 92
Glu211	A: 100	A: 100	A: 99	A: 52	A: 100	A: 57	A: 88

[a] The neutral and anionic forms of chromophore are denoted by N and A, respectively. The neutral and cationic His193 is denoted by HIE and HIP, respectively. The anionic Glu144 and Glu211 are denoted by A. [b] The ONIOM-EE-optimized structure for $\mathbf{N_{trans}}$ in the off-state protein is used. [c] The ONIOM-EE-optimized structure for $\mathbf{A_{cis}}$ in the on-state protein is used. [d] The ONIOM-EE-optimized structure for $\mathbf{Z_{cis}}$ in the on-state protein is used.

divided into two parts: the geometrical effect and the electronic effect. The estimated geometrical effect can be attributed by the protein effect on the chromophore geometry, and is evaluated as the absorption or emission energy difference of the QM part optimized in the gas phase and that in the ONIOM-optimized protein. The electronic effect includes QM−MM electrostatic interaction, as well as electronic polarization of the QM part due to the MM point charges.[68] In Table 5, the difference between columns 4 and 3, $\Delta E_{QM,model} - \Delta E_{gas}$, represents the geometrical effect and the difference between columns 5 and 4, $\Delta E_{ONIOM\text{-}EE} - \Delta E_{QM,model}$, represents the electronic effect on the absorption and emission energies for $\mathbf{A_{cis}}$ as well as $\mathbf{N_{trans}}$.

Concerning the on-state $\mathbf{A_{cis}}$, using the ONIOM SAC-CI method at ONIOM(B3LYP:AMBER)-ME- and ONIOM(CASSCF:AMBER)-ME-optimized ground-state structures, the geometric effect shifts the absorption energy by 0.17−0.22 eV to the red, which is partially canceled by the blue-shift of 0.05−0.09 eV due to the electronic effect.[69] Furthermore, for the emission of $\mathbf{A_{cis}}$ calculated by the ONIOM SAC-CI and ONIOM CASPT2 methods at the ONIOM(CASSCF:AMBER)-ME-optimized excited-state geometry in the protein, the overall protein effect of red-shift by about 0.21 eV (SAC-CI) and 0.04 eV (CASPT2) is dominated by the geometrical effect. Concerning the off-state $\mathbf{N_{trans}}$, irrespective of the ONIOM(B3LYP:AMBER)- or ONIOM(CASSCF:AMBER)-optimized geometries, the ONIOM SAC-CI calculated overall protein effect of 0.37−0.45 eV red-shift for absorption has the contribution of the geometrical effect of 0.29−0.24 eV red-shift enhanced by the electronic effect of 0.08−0.21 eV red-shift. On the other hand, compared to the absorption, a very large overall protein effect of 0.64 eV red-shift for emission has the contribution of the geometrical effect of 0.12 eV red-shift significantly enhanced by the electronic effect of 0.50 eV red-shift, with the ONIOM SAC-CI method.

In summary, by means of extensive and systematic high-level QM and ONIOM calculations on the vertical absorption and emission energies in the proteins, $\mathbf{A_{cis}}$ and $\mathbf{N_{trans}}$ are supported as the dominant protonation states of the chromophore in the on- and off-states of Dronpa, respectively. The feasibility of $\mathbf{Z_{cis}}$ is not consistently supported. Our calculations also suggest that the calculated absorption and emission energies for the off-state agree with the experiment only when the protein effects are included. These results support the *cis*−*trans* isomerization of the chromophore along with the change of protonation state.[12d]

C. Protonation Probability in the Proteins. In order to further verify the protonation states of the on- and off-states of Dronpa, we have computed the protonation possibilities of the chromophore and its key nearby titratable residues by solving the linear Poisson−Boltzmann electrostatics (PBE) equation and MC simulations, as explained in the method section. These calculations provide an independent examination for the protonation states of the chromophore in the ground state. Also, these calculations allow examining the effect of the local environment around the chromophore. Preliminary DFT calculations suggested that the pK_a of one titratable site (i.e., the phenol oxygen or imidazolinone nitrogen) of the chromophore depends much on the protonation state of the other titratable site via π-conjugation. Therefore, we will calculate the pK_a of one titratable site by keeping the protonation state of the other site invariable in each protonation possibility calculation.

The calculated population for different protonation states of the chromophore and its key nearby residues, Glu144, His193, and Glu211, at pH 7 is summarized in Table 6. First, we discuss the protonation state in the off-state Dronpa with the *trans* chromophore. The PBE calculations in columns 2 and 3 of Table 6 suggest the neutral *trans* chromophore (100%), irrespective of the geometry taken from the X-ray or ONIOM-EE-optimized structure.[70] In addition, Glu144 and Glu211 are suggested to be anionic, but His193 is in a neutral form. On the other hand, in the on-state Dronpa with the *cis* chromophore (columns 4 and 5), phenol of the chromophore is calculated to be essentially

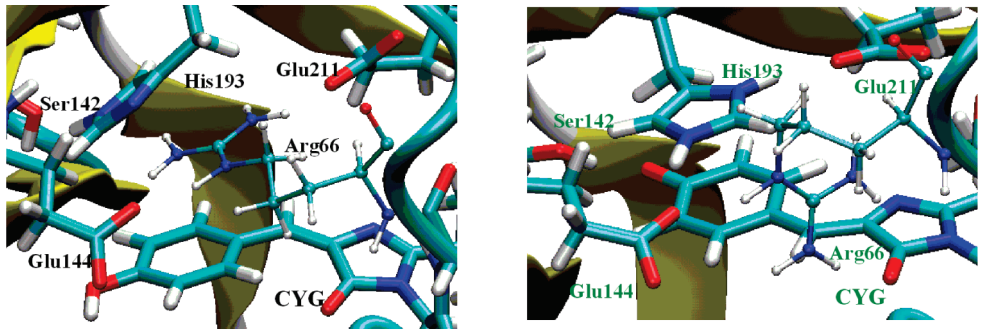


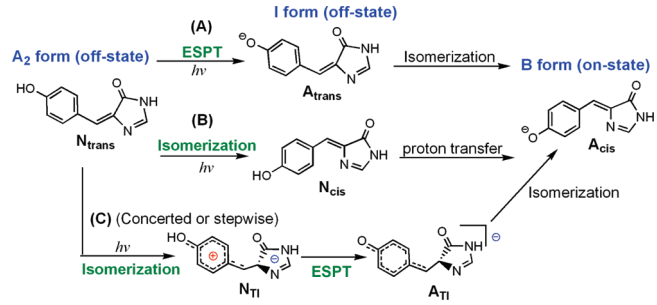
Figure 3. Immediate protein environment around the chromophores in the off-state (left) and on-state (right) of Dronpa.

in an anionic form (94−100%), Glu144 and Glu211 are anionic, and cationic His193 forms a salt bridge when the ONIOM-EE-optimized structure is used. However, the population to form the salt bridge between Glu211 and His193 is decreased to roughly 50% when the X-ray structure with a longer hydrogen-bonding distance between Glu211 and His193 was used. To study the stability of this salt bridge, we further performed the ONIOM-ME optimizations by including the side-chain of Glu211 and His193 in the QM region for the following two situations: (1) neutral His193 protonated at ε nitrogen and neutral Glu211 and (2) cationic His193 and anionic Glu211. These ONIOM calculations showed that the latter situation is energetically more favorable than the former. Therefore, in the on-state, His193 is suggested to be doubly protonated and Glu211 is deprotonated. Finally, the calculated population of $\mathbf{Z}_{cis}$ is much smaller than $\mathbf{A}_{cis}$ in the on-state protein (columns 6−8), no matter which geometry was adopted. As a result, these PBE calculations suggest that the local protein environment and *trans*−*cis* isomerization of the chromophore modulate the protonation state of the chromophore, and that the feasibility of $\mathbf{Z}_{cis}$ is reduced further.

As depicted in Figure 3, the immediate local protein environment around the chromophore is different between the on-state and off-state Dronpa. For instance, in the off-state protein, anionic Glu144 which is close to the phenol oxygen of the chromophore should disfavor the anionic form of the chromophore. Arg66 forms a salt bridge with Glu144 and Glu211. On the other hand, in the on-state protein, Ser142 forms a hydrogen bond with the phenol oxygen of the chromophore and cationic His193 forms a salt bridge with Glu144 and Glu211, although Glu144 becomes slightly far from the chromophore and Arg66 interacts with the carbonyl oxygen of the imidazole ring. His193 obviously changes the protonation state between the on- and off-state proteins, and might be regarded as a potential proton acceptor. Such different local protein environment could finely tune the protonation state of the chromophore. To qualitatively estimate the effect of the electrostatic interaction from these nearby residues (Glu144, Ser142, His193, and Glu211) on the protonation state in the on-state Dronpa, we additionally performed the PBE calculations by turning off the charges on the side-chain of one of these residues. The population of $\mathbf{A}_{cis}$ changed very little by eliminating the charges of Glu144, Glu211, or His193 (∼93−100%). However, the population of $\mathbf{A}_{cis}$ is reduced from 100% to about 46−66% by excluding the charges of Ser142, in agreement with the experimental results of a very recent mutation study, where mutation of Ser142 by Ala, Asp, Cys, or Gly was shown to afford the neutral chromophore.[13] The importance of Ser142 is in stabilizing the phenoxide anion through the hydrogen bond.[71]

Summarizing, the protonation state in the on- and off-state Dronpa was further examined by the Poisson−Boltzmann

SCHEME 4: Possible Reaction Mechanisms in Dronpa



electrostatics calculations, which support $\mathbf{A}_{cis}$ and $\mathbf{N}_{trans}$ as the on- and off-state of Dronpa, respectively, and also eliminate the possibility of $\mathbf{Z}_{cis}$. Our PBE calculations also support that the local protein environment around the chromophore, especially, Ser142, influences the protonation state of the chromophore, and thus could dictate the fluorescent "on" and nonfluorescent "off" states.[12d,13] Therefore, the proposed mechanism involving the *cis*−*trans* isomerization of the chromophore followed by a change of the protonation state of the chromophore influenced by the local protein environment is strongly suggested as one of the possible mechanisms by our calculations.[12d]

D. Reaction Mechanism. The reaction mechanism of the reversibly photochromic properties of Dronpa is not clear and still under debate. The elusive reaction mechanism of the photochromic fluorescent proteins is suggested to be controlled by several factors: protonation states of the chromophore, conformations of the chromophore, nonplanarity or flexibility of the chromophore, or intersystem crossing.[2b,3,9,12d,13,14e] Miyawaki and co-workers first proposed that the reaction is initiated by excited-state proton transfer (ESPT) from the neutral chromophore in the off-state Dronpa and gives the anionic chromophore in the on-state Dronpa (analogous to mechanism A in Scheme 4).[3] The ESPT mechanism was supported by the measured KIE.[3e] By employing the X-ray crystal structures of the on- and off-states of Dronpa, our extensive calculations discussed above implicate the proposed mechanism involving isomerization of the chromophore leading to changes of the protonation state of the chromophore as a feasible pathway (mechanism B in Scheme 4).[12d] However, the barrier of the *trans*−*cis* isomerization (the first step in mechanism B), probably via a hula-twisted pathway,[72] is still unknown. Also, this proposed pathway may not explain the origin of the observed KIE in Dronpa.[3e] The mechanism simply involving excited-state proton transfer (ESPT) without consideration of the *trans* conformation of the chromophore seems to be inconsistent with the recent X-ray crystal structure of the off-state of Dronpa, which was recently questioned because the electron density for the phenol group was not well resolved.[12d,13]

Consequently, it may alternatively be interpreted as multiple conformations of the chromophore, or a substantial disorder of the phenol group in the off-state crystal structure. It should be noted that the chromophore in the off-state is in a *trans* conformation in X-ray crystal structures of other RPFP variants, mTFP0.7, asFP595, and IrisFP.[6,73]

To account for the observed KIE and conformational change of the chromophore in the off-state crystal structure, we proposed an alternative feasible mechanism C in Scheme 4. In this mechanism, the excitation of $\mathbf{N_{trans}}$, which weakens the R5 bond and strengthens the R6 bond (Figures S12−S14 in the Supporting Information), should facilitate photoisomerization along the R5 bond. Importantly, our calculations showed that a photoisomerization along the R5 bond gives a stable twisted minimum $\mathbf{N_{TI}}$.[74−76] The imidazolinone and phenol rings in $\mathbf{N_{TI}}$ bear about −0.89 and +0.49 e (i.e., twisted intramolecular charge transfer (TICT) state, see Figure S19 in the Supporting Information), respectively. Therefore, the acidity of the phenol is enhanced by photoisomerization along the R5 bond, which promotes ESPT to afford an anionic twisted intermediate ($\mathbf{A_{TI}}$) in a concerted or stepwise manner. Isomerization of $\mathbf{A_{TI}}$ eventually gives $\mathbf{A_{cis}}$. The proton acceptor around the chromophore is vital for the ESPT process in wild-type GFP.[11a,e,77] In all available X-ray crystal structures for the dark-state RPFPs, one potential proton acceptor, glutamic acid residue (i.e., Glu144 in Dronpa, Glu148 in mTFP0.7, Glu145 in asFP595, and Glu144 in IrisFP)[6,12d,73] which seems to be analogous to Glu222 in *wt*GFP, forms a hydrogen-bond network with the phenol oxygen of the chromophore via one water molecule and therefore may be critical in the ESPT step.

As for the reaction mechanism of the on-to-off (photobleaching) process, Miyawaki and co-workers proposed an intriguing possibility that the protonation follows an intersystem crossing.[13] The minor populated $\mathbf{N_{cis}}$ may then undergo *cis*-to-*trans* isomerization to give $\mathbf{N_{trans}}$. The *cis*-to-*trans* isomerization in GFP via the neutral, rather than anionic, *cis* chromophore was also suggested by the previous semiempirical calculations.[16a] In addition, the recent CASSCF(6e,6o)/3-21G//OPLS molecular dynamics simulations on asFP595 suggested that the neutral chromophores are responsible for the isomerization, and have a higher probability for the *cis*-to-*trans* isomerization than the *trans*-to-*cis* isomerization.[17b] Moreover, much weaker $\pi-\pi$ interaction of $\mathbf{N_{cis}}$ with protonated (or possibly neutral) His193 than $\mathbf{A_{cis}}$ could facilitate the *cis*-to-*trans* isomerization, as that interaction was shown to be important for the on-state in the recent mutation studies.[13]

IV. Conclusions

The nature of the proposed "on" state and "off" state of the reversibly photochromic fluorescent protein Dronpa was extensively examined by QM and ONIOM(QM:MM) calculations. The vertical absorption and emission energies of the chromophore with several possible protonation states and two conformations in the gas phase and proteins were studied by several high-level QM (CASPT2 and SAC-CI) and ONIOM-(QM:MM) calculations. The ONIOM calculations support the neutral form as the off-state (*trans* conformation) and the anionic form as the on-state (*cis* conformation) of Dronpa. These protonation states are further enforced by the protonation probability calculations via Poisson−Boltzmann electrostatics and Monte Carlo sampling. Moreover, the local protein environments were found to influence protonation states of the chromophore. Our calculations also rule out the possibility of the zwitterionic *cis* form as the on-state of Dronpa, though the zwitterionic form was proposed to be involved in another reversibly photochromic fluorescent protein asFP595.

Possibilities of different reaction mechanisms are also discussed. Especially, we proposed an alternative mechanism C in Scheme 4, involving the photoisomerization followed by excited-state proton transfer in a concerted or stepwise manner, and finally isomerization to account for the available experimental observations (KIE and crystal structures). More studies are required to resolve the mysterious mechanism of the reversible photochromic properties in the new fluorescent proteins. Our theoretical studies on the reaction mechanism of Dronpa via photoisomerization and excited-state proton transfer are in progress and will be reported in due course.

Acknowledgment. We are grateful to Prof. Thom Vreven and Prof. Masahiro Ehara for helpful discussions on ONIOM-EE and SAC-CI calculations, respectively. We also thank Prof. Donald Bashford for providing the program MEAD and suggestion. This work is in part supported by Japan Science and Technology Agency (JST) with a Core Research for Evolutional Science and Technology (CREST) grant in the Area of High Performance Computing for Multiscale and Multiphysics Phenomena. L.W.C. acknowledges the Fukui Institute Fellowship. The computational resource at Research Center of Computer Science (RCCS) at the Institute for Molecular Science (IMS) is also acknowledged.

Supporting Information Available: Complete citation of refs 21, 23, and 33, Figures S1−S20, Tables S1−S5, Cartesian coordinates and absolute energies in the gas phase. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) (a) Conn, P. M., Ed. *Methods Enzymol.*; Academic Press: San Diego, CA, 1999; Vol. 302. (b) Zimmer, M. *Chem. Rev.* **2002**, *102*, 759. (c) Remington, S. J. *Curr. Opin. Struct. Biol.* **2006**, *16*, 714. (d) Tsien, R. Y. *Annu. Rev. Biochem.* **1998**, *67*, 509. (e) *Green Fluorescent Protein: Properties, Applications and Protocols*; Chalfie, M., Kain, S. R., Eds.; Wiley-Liss: Hoboken, NJ, 2005. (f) Sullivan, K. F., Ed. *Methods in Cell Biology*, 2nd ed.; Academic Press: London, 2008; Vol. 85. (g) Lippincott-Schwartz, J.; Altan-Bonnet, N.; Patterson, G. H. *Nat. Cell Biol.* **2003**, *5*, S7. (h) Miyawaki, A.; Sawano, A.; Kogure, T. *Nat. Cell Biol.* **2003**, *5*, S1. (i) Chudakov, D. M.; Lukyanov, S.; Lukyanov, K. A. *Trends Biotechnol.* **2005**, *23*, 605. (j) Zhang, J.; Campbell, R. E.; Ting, A. Y.; Tsien, R. Y. *Nat. Rev. Mol. Cell Biol.* **2002**, *3*, 906.

(2) (a) Shaner, N. C.; Patterson, G. H.; Davidson, M. W. *J. Cell Sci.* **2007**, *120*, 4247. (b) Henderson, J. N.; Remington, S. *J. Physiol.* **2006**, *21*, 161. (c) Lukyanov1, K. A.; Chudakov1, D. M.; Lukyanov, S.; Verkhusha, V. V. *Nat. Rev. Mol. Cell Biol.* **2005**, *6*, 885. (d) Lippincott-Schwartz, J.; Patterson, G. H. *Science* **2003**, *300*, 87.

(3) (a) Ando, R.; Mizuno, H.; Miyawaki, A. *Science* **2004**, *306*, 1370. (b) Habuchi, S.; Ando, R.; Dedecker, P.; Verheijen, W.; Mizuno, H.; Miyawaki, A.; Hofkens, J. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 9511. (c) Dedecker, P.; Hotta, J.; Ando, R.; Miyawaki, A.; Engelborghs, Y.; Hofkens, J. *Biophys. J.* **2006**, *91*, L45. (d) Habuchi, S.; Dedecker, P.; Hotta, J. I.; Flors, C.; Ando, R.; Mizuno, H.; Miyawaki, A.; Hofkens, J. *Photochem. Photobiol. Sci.* **2006**, *5*, 567. (e) Fron, E.; Flors, C.; Schweitzer, G.; Habuchi, S.; Ando, R.; De Schryver, F. C.; Miyawaki, A.; Hofkens, J. *J. Am. Chem. Soc.* **2007**, *129*, 4870.

(4) (a) Ando, R.; Hama, H.; Yamamoto-Hino, M.; Mizuno, H.; Miyawaki, A. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 12651. (b) Wiedenmann, J.; Ivanchenko, S.; Oswald, F.; Schmitt, F.; Röcker, C.; Salih, A.; Spindler, K. D.; Nienhaus, G. U. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 15905. (c) Nienhaus, G. U.; Wiedenmann, J.; Nar, H. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 9156. (d) Gurskaya, N. G.; Verkhusha, V. V.; Shcheglov, A. S.; Staroverov, D. B.; Chepurnykh, T. V.; Fradkov, A. F.; Lukyanov, S.; Lukyanov, K. A. *Nat. Biotechnol.* **2006**, *24*, 461. (e) Patterson, G. H.; Lippincott-Schwartz, J. *Science* **2002**, *297*, 1873.

(5) (a) Chudakov, D. M.; Belousov, V. V.; Zaraisky, A. G.; Novoselov, V. V.; Staroverov, D. B.; Zorov, D. B.; Lukyanov, S.; Lukyanov, K. A. *Nat. Biotechnol.* **2003**, *21*, 191. (b) Chudakov, D. M.; Feofanov, A. V.; Mudrik, N. N.; Lukyanov, S.; Lukyanov, K. A. *J. Biol. Chem.* **2003**, *278*, 7215. (c) Chudakov, D. M.; Verkhusha, V. V.; Staroverov, D. B.; Souslova,

E. A.; Lukyanov, S.; Lukyanov, K. A. *Nat. Biotechnol.* **2004**, *22*, 1435. (d) Miyawaki, A. *Nat. Biotechnol.* **2004**, *22*, 1374. (e) Verkhusha, V. V.; Lukyanov, K. A. *Nat. Biotechnol.* **2004**, *22*, 289. (f) Lukyanov, K. A.; Fradkov, A. F.; Gurskaya, N. G.; Matz, M. V.; Labas, Y. A.; Savitskyi, A. P.; Markelov, M. L.; Zaraisky, A. G.; Zhao, X.-N.; Fang, Y.; Tan, W.-Y.; Lukyanov, S. A. *J. Biol. Chem.* **2000**, *275*, 25879.

(6) (a) Andresen, M.; Wahl, M. C.; Stiel, A. C.; Gräter, F.; Schäfer, L. V.; Trowitzsch, S.; Weber, G.; Eggeling, C.; Grubmüller, H.; Hell, S. W.; Jakobs, S. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 13070. (b) Henderson, N. J.; Ai, H.-W.; Campbell, R. E.; Remington, S. J. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 6672. (c) Hofmann, M.; Eggeling, C.; Jakobs, S.; Hell, S. W. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 17565.

(7) Wilmann, P. G.; Petersen, J.; Devenish, R. J.; Prescott, M.; Rossjohn, J. *J. Biol. Chem.* **2005**, *280*, 2401.

(8) Sauer, M. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 9433.

(9) (a) Ando, R.; Flors, C.; Mizuno, H.; Hofkens, J.; Miyawaki, A. *Biophys. J.* **2007**, *92*, L97. (b) Flors, C.; Hotta, J.-I.; Uji-i, H.; Dedecker, P.; Ando, R.; Mizuno, H.; Miyawaki, A.; Hofkens, J. *J. Am. Chem. Soc.* **2007**, *129*, 13970.

(10) During the preparation of this manuscript, bsDronpa and Padron, mutants of Dronpa, have been reported: Andresen, M.; Stiel, A. C.; Fölling, J.; Wenzel, D.; Schönle, A.; Egner, A.; Eggeling, C.; Hell, S. W.; Jakobs, S. *Nat. Biotechnol.* **2008**, *26*, 1035.

(11) (a) Brejc, K.; Sixma, T. K.; Kitts, P. A.; Kain, S. R.; Tsien, R. Y.; Ormoe, M.; Remington, S. J. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 2306. (b) Heim, R.; Prasher, D. C.; Tsien, R. Y. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 12501. (c) Palm, G. J.; Zdanov, A.; Gaitanaris, G. A.; Stauber, R.; Pavlakis, G. N.; Wlodawer, A. *Nat. Struct. Biol.* **1997**, *4*, 361. (d) Chattoraj, M.; King, B. A.; Bublitz, G. U.; Boxer, S. G. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 8362. (e) Stoner-Ma, D.; Jaye, A. A.; Matousek, P.; Towrie, M.; Meech, S. R.; Tonge, P. J. *J. Am. Chem. Soc.* **2005**, *127*, 2864.

(12) (a) Wilmann, P. G.; Turcic, K.; Battad, J. M.; Wilce, M. C. J.; Devenish, R. J.; Prescott, M.; Rossjohn, J. *J. Mol. Biol.* **2006**, *364*, 213. (b) Stiel, A. C.; Trowitzsch, S.; Weber, G.; Andresen, M.; Eggeling, C.; Hell, S. W.; Jakobs, S.; Wahl, M. C. *Biochem. J.* **2007**, *402*, 35. (c) Nam, K.-H.; Kwon, O. Y.; Sugiyama, K.; Lee, W.-H.; Ki, Y. K.; Song, H. K.; Kim, E. E.; Park, S.-Y.; Jeon, H.; Hwang, K. S. *Biochem. Biophys. Res. Commun.* **2007**, *354*, 962. (d) Andresen, M.; Stiel, A. C.; Trowitzsch, S.; Weber, G.; Eggeling, C.; Wahl, M. C.; Hell, S. W.; Jakobs, S. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 13005.

(13) During the preparation of this manuscript, one additional X-ray structure of the on-state was obtained: Mizuno, H.; Kumar, M. T.; Wälchli, M.; Kikuchi, A.; Fukano, T.; Ando, R.; Jeyakanthan, J.; Taka, J.; Shiro, Y.; Ikura, M.; Miyawaki, A. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 9927.

(14) (a) Petersen, J.; Wilmann, P. G.; Beddoe, T.; Oakley, A. J.; Devenish, R. J.; Prescott, M.; Rossjohn, J. *J. Biol. Chem.* **2003**, *278*, 44626. (b) Prescott, M.; Ling, M.; Beddoe, T.; Oakley, A. J.; Dove, S.; Hoegh-Guldberg, O.; Devenish, R. J.; Rossjohn, J. *Structure* **2003**, *11*, 275. (c) Wilmann, P. G.; Petersen, J.; Pettikiriarachchi, A.; Buckle, A. M.; Smith, S. C.; Olsen, S.; Perugini, M. A.; Devenish, R. J.; Prescott, M.; Rossjohn, J. *J. Mol. Biol.* **2005**, *349*, 223. (d) Loos, D. C.; Habuchi, S.; Flors, C.; Hotta, J.-I.; Wiedenmann, J.; Nienhaus, G. U.; Hofkens, J. *J. Am. Chem. Soc.* **2006**, *128*, 6270. (e) Quillin, M. L.; Anstrom, D. M.; Shu, X.; O'Leary, S.; Kallio, K.; Chuadakov, D. M.; Remington, S. J. *Biochemistry* **2005**, *44*, 5774. (f) Nienhaus, K.; Nar, H.; Heilker, R.; Wiedenmann, J.; Nienhaus, G. U. *J. Am. Chem. Soc.* **2008**, *130*, 12578.

(15) (a) Martin, M. E.; Negri, F.; Olivucci, M. *J. Am. Chem. Soc.* **2004**, *126*, 5452. (b) Altoe, P.; Bernardi, F.; Garavelli, M.; Orlandi, G.; Negri, F. *J. Am. Chem. Soc.* **2005**, *127*, 3952. (c) Sinicropi, A.; Andruniow, T.; Ferre, N.; Basosi, R.; Olivucci, M. *J. Am. Chem. Soc.* **2005**, *127*, 11534. (d) Lopez, X.; Marques, M. A. L.; Castro, A.; Rubio, A. *J. Am. Chem. Soc.* **2005**, *127*, 12329. (e) Vendrell, O.; Gelabert, R.; Moreno, M.; Lluch, J. M. *J. Am. Chem. Soc.* **2006**, *128*, 3564. (f) Olsen, S.; Smith, S. C. *J. Am. Chem. Soc.* **2007**, *129*, 2054. (g) Olsen, S.; Smith, S. C. *J. Am. Chem. Soc.* **2008**, *130*, 8677. (h) Toniolo, A.; Olsen, S.; Manohar, L.; Martinez, T. J. *Faraday Discuss.* **2004**, *127*, 149. (i) Hasegawa, J.-Y.; Fujimoto, K.; Swerts, B.; Miyahara, T.; Nakatsuji, H. *J. Comput. Chem.* **2007**, *28*, 2443. (j) Das, A. K.; Hasegawa, J.; Miyahara, T.; Ehara, M.; Nakatsuji, H. *J. Comput. Chem.* **2003**, *24*, 1421. (k) Demachy, I.; Ridard, J.; Laguitton-Pasquier, H.; Durnerin, E.; Vallverdu, G.; Archirel, P.; Levy, B. *J. Phys. Chem. B* **2005**, *109*, 24121. (l) Nifosí, R.; Amat, P.; Tozzini, V. *J. Comput. Chem.* **2007**, *28*, 2366. (m) Bravaya, K. B.; Bochenkova, A. V.; Granovsky, A. A.; Savitsky, A. P.; Nemukhin, A. V. *J. Phys. Chem. A* **2008**, *112*, 8804. (n) Nemukhin, A. V.; Topol, I. A.; Burt, S. K. *J. Chem. Theory Comput.* **2006**, *2*, 292. (o) Amat, P.; Granucci, G.; Buda, F.; Persico, M.; Tozzini, V. *J. Phys. Chem. B* **2006**, *110*, 9348. (p) Voityuk, A. A.; Kummer, A. D.; Michel-Beyerle, M.-E.; Rosch, N. *Chem. Phys.* **2001**, *269*, 83.

(16) (a) Weber, W.; Helms, V.; Mccammon, J. A.; Langhoffi, P. W. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 6177. (b) Helms, V. *Curr. Opin. Struct. Biol.* **2002**, *12*, 169.

(17) (a) Schäfer, L. V.; Groenhof, G.; Klingen, A. R.; Ullmann, G. M.; Boggio-Pasqua, M.; Robb, M. A.; Grubmüller, H. *Angew. Chem., Int. Ed.* **2007**, *46*, 530. (b) Schäfer, L. V.; Groenhof, G.; Boggio-Pasqua, M.; Robb, M. A.; Grubmüller, H. *PLoS Comput. Biol.* **2008**, *4*, 1.

(18) For classical molecular dynamic simulations on Dronpa and Dronpa-2, see: (a) Moors, S. L. C.; Michielssens, S.; Flors, C.; Dedecker, P.; Hofkens, J.; Ceulemans, A. *J. Chem. Theory Comput.* **2008**, *4*, 1012. (b) Asselberghs, I.; Flors, C.; Ferrighi, L.; Botek, E.; Champagne, B.; Mizuno, H.; Ando, R.; Miyawaki, A.; Hofkens, J.; Van der Auweraer, M.; Clays, K. *J. Am. Chem. Soc.* **2008**, *130*, 15713.

(19) (a) Hooft, R. W. W.; Vriend, G.; Sander, C.; Abola, E. E. *Nature* **1996**, *381*, 272. (b) Lovell, S. C.; Davis, I. W.; Arendall III, W. B.; de Bakker, P. I. W.; Word, J. M.; Prisant, M. G.; Richardson, J. S.; Richardson, D. C. *Proteins: Struct., Funct., Genet.* **2003**, *50*, 437.

(20) (a) Dolinsky, T. J.; Nielsen, J. E.; McCammon, J. A.; Baker, N. A. *Nucleic Acids Res.* **2004**, *32*, W665. (b) Li, H.; Robertson, A. D.; Jensen, J. H. *Proteins: Struct., Funct., Bioinf.* **2005**, *61*, 704.

(21) Case, D. A.; et al. *AMBER 9*; University of California: San Francisco, CA, 2006.

(22) (a) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179. (b) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Comput. Phys.* **1983**, *79*, 926.

(23) (a) Frisch, M. J.; et al. *Gaussian Development Version*, revisions G.01 and G.03; Gaussian, Inc.: Wallingford, CT, 2004.

(24) (a) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089. (b) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577.

(25) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327.

(26) (a) Pastor, R. W.; Brooks, B. R.; Szabo, A. *Mol. Phys.* **1988**, *65*, 1409. (b) Loncharich, R. J.; Brooks, B. R.; Pastor, R. W. *Biopolymers* **1992**, *32*, 523. (c) Izaguirre, J. A.; Catarello, D. P.; Wozniak, J. M.; Skeel, R. D. *J. Chem. Phys.* **2001**, *114*, 2090.

(27) (a) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648. (b) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.

(28) (a) Werner, H.-J.; Knowles, P. J. *J. Chem. Phys.* **1985**, *82*, 5053. (b) Knowles, P. J.; Werner, H.-J. *Chem. Phys. Lett.* **1985**, *115*, 259.

(29) (a) Stratmann, R. E.; Scuseria, G. E.; Frisch, M. J. *J. Chem. Phys.* **1998**, *109*, 8218. (b) Bauernschmitt, R.; Ahlrichs, R. *Chem. Phys. Lett.* **1996**, *256*, 454. (c) Celani, P.; Werner, H.-J. *J. Chem. Phys.* **2000**, *112*, 5546. The SAC-CI calculations using singles and doubles linked excitation operators were performed: (d) Nakatsuji, H. *Chem. Phys. Lett.* **1978**, *59*, 362. (e) Nakatsuji, H. *Chem. Phys. Lett.* **1989**, *67*, 329. (f) Nakatsuji, H. *Chem. Phys. Lett.* **1989**, *67*, 334.

(30) (a) The calculated absorption and emission energies by the SAC-CI method were found to be slightly changed in some cases by the number of states for the perturbation selection. The SAC-CI calculations by solving five states were used in all calculations for the consistency. (b) The SAC-CI method is theoretically similar to the EOM-CCSD and CC linear response (LR) methods. Therefore, the SAC-CI method should be more reliable than TD-DFT, as shown in GFP (ref 15j). In addition, recent SAC-CI work gave a good agreement on other fluorescent proteins, e.g., BFP and CFP (ref 15i). The detailed discussion and comparison of SAC-CI and the other coupled-cluster methods can be found in the following reviews: (c) Ehara, M.; Hasegawa, J.; Nakatsuji, H. In *Theory and Applications of Computational Chemistry: The First 40 Years*; Dykstra, C. E., Frenking, G., Kim, K. S., Scuseria, G. E., Eds.; Elsevier: Oxford, U.K., 2005; pp 1099−1141. (d) Bartlett, R. J. In *Theory and Applications of Computational Chemistry: The First 40 Years*; Dykstra, C. E., Frenking, G., Kim, K. S., Scuseria, G. E., Eds.; Elsevier: Oxford, U.K., 2005; pp 1191−1221.

(31) Orbital rotation of the phenolic oxygen lone-pair with the amidic lone-pair occurs in few calculations, due to double occupation of the former one.

(32) (a) Celani, P.; Werner, H.-J. *J. Chem. Phys.* **2000**, *112*, 5546. (b) Roos, B. O.; Andersson, K. *Chem. Phys. Lett.* **1995**, *245*, 215. (c) Andersson, K.; Malmqvist, P.-Å.; Roos, B. O. *J. Chem. Phys.* **1992**, *96*, 1218.

(33) Werner, H.-J.; Knowles, P. J.; Lindh, R.; Manby, F. R.; Schütz, M.; et al. *MOLPRO, version 2006.1, a package of ab initio programs* (see http://www.molpro.net).

(34) Merrick, J. P.; Moran, D.; Radom, L. *J. Phys. Chem. A* **2007**, *111*, 11683.

(35) (a) Bravaya, K. B.; Bochenkova, A. V.; Granovskii, A. A.; Nemukhin, A. V. *Russ. J. Phys. Chem. B* **2008**, *2*, 671. The calculated absorption for the neutral form of the model GFP chromophore (HBIA) is 3.61, 3.71, and 3.11 eV by the MRMP2, MCQDPT2, and aug-MCQDPT2 methods, respectively. During the reviewing process, one theoretical paper for the GFP model was reported: (b) Filippi, C.; Zaccheddu, M.; Buda, F. *J. Chem. Theory Comput.* **2009**, *5*, 2074.

(36) Due to the high computational cost for SA4-CASSCF optimizations, we mostly used a slightly smaller active space (12e,12o) shown in Figures S7 and S8 in the Supporting Information, which gives similar structures and absorption/emission energies obtained by the larger active space (14e,13o) (Figure S14 and Table S3 in the Supporting Information).

(37) (a) Karlström, G.; Lindh, R.; Malmqvist, P.-Å.; Roos, B. O.; Ryde, U.; Veryazov, V.; Widmark, P.-O.; Cossi, M.; Schimmelpfennig, B.; Neograd y, P.; Seijo, L. *Comput. Mater. Sci.* **2003**, *28*, 222. (b) Forsberg, N.; Malmqvist, P.-Å. *Chem. Phys. Lett.* **1997**, *274*, 196.

(38) Ghigo, G.; Roos, B. O.; Malmqvist, P.-Å. *Chem. Phys. Lett.* **2004**, *396*, 142.

(39) (a) Maseras, F.; Morokuma, K. *J. Comput. Chem.* **1995**, *16*, 1170. (b) Humbel, S.; Sieber, S.; Morokuma, K. *J. Chem. Phys.* **1996**, *105*, 1959. (c) Matsubara, T.; Sieber, S.; Morokuma, K. *Int. J. Quantum Chem.* **1996**, *60*, 1101. (d) Svensson, M.; Humbel, S.; Froese, R. D. J.; Matsubara, T.; Sieber, S.; Morokuma, K. *J. Phys. Chem.* **1996**, *100*, 19357. (e) Svensson, M.; Humbel, S.; Morokuma, K. *J. Chem. Phys.* **1996**, *105*, 3654. (f) Dapprich, S.; Komáromi, I.; Byun, S.; Morokuma, K.; Frisch, M. J. *THEOCHEM* **1999**, *461*, 1. (g) Vreven, T.; Morokuma, K. *J. Comput. Chem.* **2000**, *21*, 1419. (h) Vreven, T.; Frisch, M. J.; Kudin, K. N.; Schlegel, H. B.; Morokuma, K. *Mol. Phys.* **2006**, *104*, 701.

(40) For recent ONIOM studies on photobiological systems, see: (a) Vreven, T.; Morokuma, K. *Theor. Chem. Acc.* **2003**, *109*, 125. (b) Hall, K. F.; Vreven, T.; Frisch, M. J.; Bearpark, M. J. *J. Mol. Biol.* **2008**, *383*, 106. (c) Altun, A.; Yokoyama, S.; Morokuma, K. *J. Phys. Chem. B* **2008**, *112*, 6814. (d) Bearpark, M. J.; Larkin, S. M.; Vreven, T. *J. Phys. Chem. A* **2008**, *112*, 7286. (e) Altun, A.; Yokoyama, S.; Morokuma, K. *J. Phys. Chem. B* **2008**, *112*, 16883.

(41) Altun, A.; Shaik, S.; Thiel, W. *J. Comput. Chem.* **2006**, *27*, 1324.

(42) Vreven, T.; Byun, K. S.; Komáromi, I.; Dapprich, S.; Montgomery, J. A., Jr.; Morokuma, K.; Frisch, M. J. *J. Chem. Theory Comput.* **2006**, *2*, 815.

(43) The molecular graph was created by PyMOL: DeLano, W. L. The PyMOL Molecular Graphics System (2002) on the World Wide Web (http://www.pymol.org).

(44) (a) Bashford, D.; Karplus, M. *Biochemistry* **1990**, *29*, 10219. (b) Bashford, D.; Gerwert, K. *J. Mol. Biol.* **1992**, *224*, 473. (c) Bashford, D. In *Scientific Computing in Object-Oriented Parallel Environments*; Ishikawa, Y., Oldehoeft, R. R.; Reynders, J. V. W., Tholburn, M., Eds.; Springer: Berlin, 1997; pp 233−240.

(45) Rabenstein, B. Karlsberg online manual 1999, http://agknapp.chemie.fu-berlin.de/karlsberg.

(46) (a) Sitkoff, D.; Sharp, K. A.; Honig, B. *J. Phys. Chem.* **1994**, *98*, 1978. (b) Bondi Radii were used to test a few cases and led to similar results (not shown): Bondi, A. *J. Phys. Chem.* **1964**, *68*, 441.

(47) Ishikita, H.; Knapp, E.-W. *J. Am. Chem. Soc.* **2007**, *129*, 1210, and references therein.

(48) (a) Nozaki, Y.; Tanford, C. *Methods Enzymol.* **1967**, *11*, 715. (b) Tanokura, M. *Biochim. Biophys. Acta* **1983**, *742*, 576.

(49) Liptak, M. D.; Shields, G. C. *J. Am. Chem. Soc.* **2001**, *123*, 7314.

(50) (a) Barone, V.; Cossi, M. *J. Phys. Chem. A* **1998**, *102*, 1995. (b) Cossi, M.; Scalmani, G.; Rega, N.; Barone, V. *J. Comput. Chem.* **2003**, *24*, 669.

(51) The measured pK_a of the imidazole nitrogen of *p*-HBDI in MeOH/H_2O (1:1) is about 2.36: Dong, J.; Solntsev, K. M.; Tolbert, L. M. *J. Am. Chem. Soc.* **2006**, *128*, 12038.

(52) Bell, A. F.; He, X.; Wachter, R. M.; Tonge, P. J. *Biochemistry* **2000**, *39*, 4423.

(53) The effect of the basis sets was found to be minor in the GFP (ref 15a).

(54) The chromophore becomes nonplanar, and the nitrogen center is rather pyramidal in $\mathbf{Z}_{cis}$, $\mathbf{Z}_{trans}$, and $\mathbf{C}_{cis}$, partly due to the steric repulsion of the NH bond or CO bond of the imidazolinone ring with the HC bond of the phenol ring (Figure S13 in the Supporting Information).

(55) For recent studies on *o*-HBDI and *m*-HBDI, see:(a) Chen, K.-Y.; Cheng, Y.-M.; Lai, C.-H.; Hsu, C.-C.; Ho, M.-L.; Lee, G.-H.; Chou, P.-T. *J. Am. Chem. Soc.* **2007**, *129*, 4534. (b) Dong, J.; Solntsev, K. M.; Poizat, O.; Tolbert, L. M. *J. Am. Chem. Soc.* **2007**, *129*, 10084.

(56) The absorption energy (3.12 eV) for the neutral GFP analogue chromophore was estimated from the absorption of the neutral analogue chromophore attached with an ammonium group and correction of the presence of the charged NH_3 group by the TD-B3LYP/6-311++G*//MP3 calculations: (a) Lammich, L.; Petersen, M. Å; Nielsen, M. B.; Andersen, L. H. *Biophys. J.* **2007**, *92*, 201. (b) Nielsen, S. B.; Lapierre, A.; Andersen, J. U.; Pedersen, U. V.; Tomita, S.; Andersen., L. H. *Phys. Rev. Lett.* **2001**, *87*, 228102. (c) Andersen, L. H.; Lapierre, A.; Nielsen, S. B.; Nielsen, I. B.; Pedersen, S. U.; Pedersen, U. V.; Tomita, S. *Eur. Phys. J. D* **2002**, *20*, 597.

(57) (a) Serrano-Andrés, L.; Merchán, M. *THEOCHEM* **2005**, *729*, 99. (b) Dreuw, A.; Head-Gordon, M. *J. Am. Chem. Soc.* **2004**, *126*, 4007. (c) Dreuw, A.; Head-Gordon, M. *Chem. Rev.* **2005**, *105*, 4009, and references therein.

(58) TD-DFT was reported to give good agreement of the computed absorption spectra with the experiments, when the protein was optimized by the AM1/MM level and then the QM part was taken for the TD-DFT (local density approximation (LDA)) calculations: Marques, M. A. L.; López, X.; Varsano, D.; Castro, A.; Rubio, A. *Phys. Rev. Lett.* **2003**, *90*, 258101. We found that the calculated absorption energies of $\mathbf{A}_{cis}$ (2.94 eV) and $\mathbf{N}_{cis}$ (3.24 eV) by local spin density approximation (LSDA) calculations are similar to the B3LYP calculations, when the B3LYP-optimized structures are used.

(59) In the recent theoretical study on asFP595 (ref 17a), the TDDFT method was also found to lead to large error for the anion form: "The TDDFT absorption bands of the zwitterion and the anion are too much blue-shifted, and none of the absorption bands is close to the measured absorption at 2.18eV." "The reason for the failure of TDDFT to quantitatively predict the excitation energies of the zwitterion and the anion is that the $S_0 \rightarrow S_1$ transitions involve intramolecular charge transfer from the phenyl to the imidazolinone ring, which is known to be usually rather poorly described by TDDFT."

(60) The absorption energy for $\mathbf{C}_{cis}$ was found to be overestimated by the CASPT2(14e,13o) calculation, compared to that for $\mathbf{C}_{cis}$ by the CASPT2(12e,12o) calculations and for $\mathbf{C}_{trans}$ by the CASPT2(14e,13o) and CASPT2(12e,12o) calculations (Table 2 and Table S1 in the Supporting Information). $\mathbf{C}_{cis}$ by the CASPT2(14e,13o) calculation is presumably subject to intruder state.

(61) Rydberg orbitals were suggested to have some effect on the conjugated system containing heteroatoms (e.g., indole) by one of the referees, as studied by Roos and co-workers: Serrano-Andrés, L.; Roos, B. O. *J. Am. Chem. Soc.* **1996**, *118*, 185. Unfortunately, the computational cost is prohibitively high to further include Rydberg orbitals (3s, 3p, and 3d orbitals) in the current systems. In addition, a good agreement between our TD-B3LYP calculation and the experiment for the neutral form suggests that the effect of the Rydberg state should not be important in the lowest electronic excitation.

(62) As displayed in Figure 2A, the chromophore for $\mathbf{A}_{cis}$ in the on-state protein shows a maximum deviation of 0.04 Å between the EE and ME optimization schemes with the small QM model (MS); this occurs at R3, i.e., the C=O bond in the imidazolinone ring. Two positively charged residues, Arg66 and Arg91, slightly elongate the carbonyl group via electrostatic interaction. The ONIOM-EE optimization also gives a noncoplanar chromophore for $\mathbf{N}_{trans}$ in the off-state protein ($\phi = 13.8°$ and $\tau = 0.4°$) and a nearly coplanar chromophore for $\mathbf{A}_{cis}$ in the on-state protein ($\phi = -7.2°$ and $\tau = -176.1°$), which are essentially the same as the ONIOM-ME optimization.

(63) The calculated maximum deviation between the two QM models for $\mathbf{A}_{cis}$ is 0.01 and 0.02 Å with the ONIOM-ME and ONIOM-EE schemes, respectively, and that for $\mathbf{N}_{trans}$ is 0.01 Å with both schemes (Figure S17 in the Supporting Information). Moreover, the nonplanarity of the chromophore for $\mathbf{N}_{trans}$ ($\phi = 17.7°$ (ME) and 15.7° (EE)) and $\mathbf{A}_{cis}$ ($\phi = -7.1°$ (ME and EE)) in the model ML1 is also similar to the model MS.

(64) See section II.B for the details.

(65) Moreover, no significant difference of the ONIOM-EE calculated absorption energies was found, when the ONIOM(B3LYP:AMBER)-ME- and ONIOM(B3LYP:AMBER)-EE-optimized structures were used. At the ONIOM-ME optimizations, the ONIOM(SAC-CI:AMBER)-EE calculated absorption energy for $\mathbf{A}_{cis}$ in the on-state protein is 2.35 eV (MS) and 2.25 eV (ML1), and that for $\mathbf{N}_{trans}$ in the off-state protein is 3.11 eV (MS) and 3.05 eV (ML1).

(66) Absorption energies and oscillator strengths calculated by the ONIOM(TD-B3LYP:AMBER)-EE method are collected in Table S4 of the Supporting Information.

(67) The calculated absorption and emission is 3.50 and 2.69 eV, respectively, when the SS-CASPT2(12e,12o)/ANO-S level as the QM method is used (Table S5 in the Supporting Information).

(68) One of the referees suggested to study the effect of the protein dynamics on the absorption and emission energies. However, for our current extensive study involving very high-level QM calculations, the computational cost is too high for reliable statistics, and this remains to be a future problem to be studied. In addition, due to the lack of very reliable force fields for the chromophore, we fixed the chromophore to the X-ray crystal structures in all classical simulations (see section II.A).

(69) The calculated geometric and electronic effects for the ONIOM(B3LYP:AMBER)-ME-optimized $\mathbf{A}_{cis}$ are about −0.15 and 0.02 eV. The geometric and electronic effects for the ONIOM(B3LYP:AMBER)-ME-optimized $\mathbf{N}_{trans}$ are about −0.17 and −0.10 eV.

(70) One of the referees suggested to perform molecular dynamics (MD) simulations for all protonation states to equilibrate the hydrogen-bond networks. However, equilibration of a global hydrogen-bonding network may require very long equilibrium MD simulations and is beyond the scope of the present paper.

(71) The pK_a of the mKate mutant (S158A) is decreased to 5.3, compared to 6.2 for the wild-type mKate: Pletnev, S.; Shcherbo, D.; Chudakov, D. M.; Pletneva, N.; Merzlyak, E. M.; Wlodawer, A.; Dauter, Z.; Pletnev, V. *J. Biol. Chem.* **2008**, *283*, 28980.

(72) (a) Liu, R. S. H. *Acc. Chem. Res.* **2001**, *34*, 555. (b) Liu, R. S. H.; Hammond, G. S. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 11153. (c) Liu, R. S. H.; Asato, A. E. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 259.

(73) Adam, V.; Lelimousin, M.; Boehme, S.; Desfonds, G.; Nienhaus, K.; Field, M. J.; Wiedenmann, J.; McSweeney, S.; Nienhaus, G. U.; Bourgeois, D. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 18343.

(74) The twisted mimum $\mathbf{N_{TI}}$ in S_1 is also supported by our CASPT2 calculations when four states were considered.

(75) A lower-barrier photoisomerization for the imidazolinone ring of the neutral GFP has recently been reported by using the SA2-CASSCF(2e,2o)/6-31G and reparameterized semiempirical (FOMO-CI) methods: Virshup, A. M.; Punwong, C.; Pogorelov, T. V.; Lindquist, B. A.; Ko, C.; Martiìnez, T. J. *J. Phys. Chem. B* **2009**, *113*, 3280.

(76) The recent excited-state QM/MM molecular dynamic simulations on asFP595 suggested that rotation of the imidazolinone ring from the neutral chromophores preferentially proceeds to give a twisted minimum (ref 17b).

(77) Stoner-Ma, D.; Jaye, A. A.; Ronayne, K. L.; Nappa, J.; Meech, S. R.; Tonge, P. J. *J. Am. Chem. Soc.* **2008**, *130*, 1227.

JP909947C