THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

Reaction Mechanism of Photo-induced Decarboxylation of the Photoactivatable Green Fluorescent Protein: An ONIOM(QM:MM) Study

Lina Ding, Lung Wa Chung, and Keiji Morokuma

J. Phys. Chem. B, **Just Accepted Manuscript** • DOI: 10.1021/jp3112952 • Publication Date (Web): 28 Dec 2012

Reaction Mechanism of Photo-induced Decarboxylation of the Photoactivatable Green Fluorescent Protein: An ONIOM(QM:MM) Study

Lina Ding,[1] Lung Wa Chung,[1] and Keiji Morokuma[1,2,*]

[1]Fukui Institute for Fundamental Chemistry, Kyoto University, 34-4 Takano Nishihiraki-cho, Sakyo-ku, Kyoto 606-8103, Japan and [2]Department of Chemistry and Cherry L. Emerson Center for Scientific Computation, Emory University, Atlanta, GA 30322

Abstract: Photoactivatable (PA) fluorescent proteins are a new class of fluorescent proteins, in which the intensity of fluorescence is dramatically enhanced through photo-induced decarboxylation process. In the present study, the reaction mechanism of the photo-induced decarboxylation in PA-GFP was investigated by the ONIOM(QM:MM) method. The decarboxylation process starts from the first excited state (IntraCT state), and then proceeds along an InterCT state after the first crossing (or an approximate transition state). Relative to an equilibrium structure in S_0, a barrier of ~ 94 kcal/mol to reach this approximate transition state is the rate-determining step for the entire decarboxylation process. The InterCT state becomes the open-shell ground state in the product, after the subsequent crossing with a closed-shell state that holds an extra electron on the dissociated CO_2. The present study elucidated for the first time the mechanism of the photo-induced decarboxylation of PA-GFP and supports the widely accepted Kolbe pathway, which could be a common mechanism for the irreversible photo-induced decarboxylation in different fluorescent proteins.

Keywords: Kolbe pathway, charge transfer, GFP, crossing

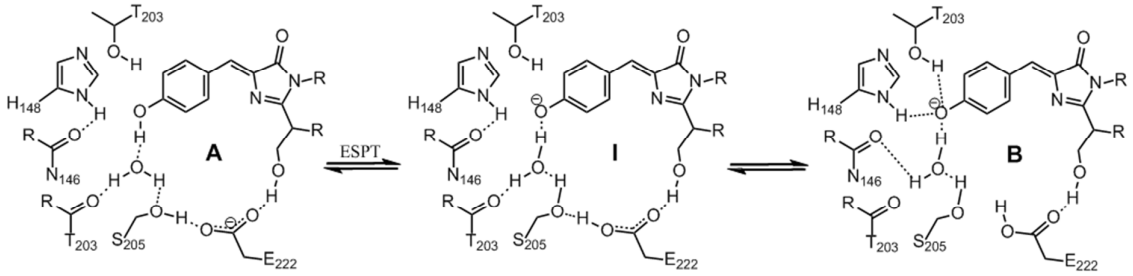
Introduction

Green fluorescent protein (GFP) and its variants have played an important role in biological imaging and analysis.[1] Photoactivatable fluorescent proteins (e.g. PA-GFP and PS-CFP),[2] irreversible photoconversion fluorescent proteins (e.g. Kaeda and EosFP)[3] and reversible photoswitching fluorescent proteins (e.g. Dronpa)[4] recently became a new class of fluorescent proteins (FPs) and provided a significant advance in fluorescent protein technology, such as super-resolution spectroscopy. Photoactivatable fluorescent proteins2 have attracted a plenty of scientific interests, since dramatic contrast by emission upon an irradiation makes them particularly popular and important as optical markers.

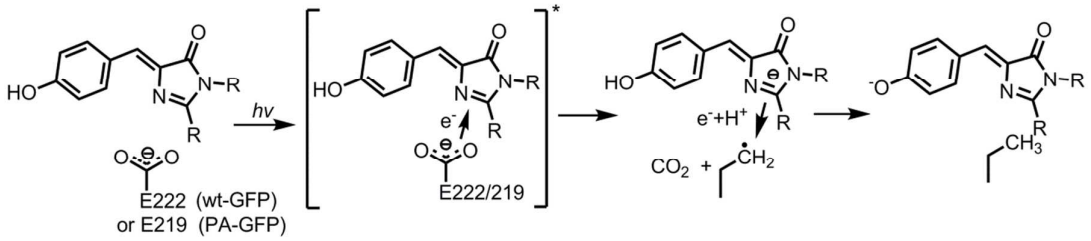
Photo-induced transformation between the neutral form (responsible to the A band) and anion form (the B band) of the GFP chromophore is essential to fluorescent emission.[5] As observed by Stark spectroscopy, the primary and key photo-physical process in the excited state is an intra-molecular charge transfer (CT) from the phenolic ring to the electron-deficient imidazolidinone ring of the chromophore.6[a,b] In general, there are two possible photo-induced transformation mechanisms (Types I and II), as observed or suggested by a number of experimental studies (such as X-ray structures and spectroscopy information). Type I mechanism is the most well-known reversible process fitting to a three-state model (A-, B-, and I-forms, see Scheme 1), in which the light-driven excited-state proton transfer (ESPT) of the A-form leads to the B-form through a transient intermediate (I-form), presumably involving a few hydrogen bonds to the nearby GLU222, a tentative proton acceptor for ESPT.[6] The A, B and I forms with distinct photo-physical properties have been identified by low-temperature hole-burning spectroscopy.6[c]

In contrast to the reversible Type I mechanism, an irreversible Type II and so-called Kolbe mechanism (Scheme 2), that accompanies the photo-induced decarboxylation of GLU222 was firstly proposed for the wild-type (WT) GFP, supported by X-ray crystallography and mass spectrometry study.[7a] The reported order of the decarboxylation rate (254 nm > 280 nm > 476 nm) showed clearly that it was excitation-wavelength dependent.7[b] Both reversible and irreversible processes are involved in the photo-transformation; in WT-GFP, the quantum yield (~0.002-0.03) for the irreversible photo-decarboxylation is much lower than that for the fluorescent quantum yield of ~ 0.8,[8]

possibly due to the competitive reversible process.1 Recently, X-ray crystal structures of native and photo-activated states of PA-GFP were also reported.[9] In addition, rate measurements for irreversible photoactivation reveal that the kinetics in PA-GFP is similar to that in WT-GFP.9



Scheme 1. Reversible Type I mechanism for the photo-induced transformation of the A- and B- forms of the chromophore via the I-form. Adapted from Brejc6[d].



Scheme 2. Irreversible Type II Kolbe mechanism for the photo-decarboxylation. Adapted from ref 7.

The Kolbe mechanism7 (Scheme 2) was proposed for the photo-induced decarboxylation in fluorescent proteins because of its similarity to the well known thermal Kolbe reaction.[10,11] The decarboxylation of GLU222 of GFP (or GLU219 of PA-GFP) takes place because of the photo-induced oxidation of GLU; upon photo-excitation one electron transfers from anionic GLU to the chromophore, followed by decarboxylation of the neutral species. Some theoretical studies on the photo-induced decarboxylation process of small organic molecules have been published, including those by Fang and co-workers.[12] Very

recently, Roy et al.[13a] reported QM/MM and MD study of the decarboxylation process of IrisFP; here however one electron was arbitrarily removed from the QM system.; Hasegawa and Krylov *et al*.[13b,c] have identified the critical CT states involved in the Kolbe mechanism which were proposed to be accessible *via* either directly one-photon absorption, or a two-step excitation via S_1.[14]

Knowledge of potential energy surfaces (PESs) of the ground and excited states involved is of great importance in understanding the reaction mechanism. In particular, it is a challenge to characterize the nature of electronic structures involved and details of the decarboxylation process. Recently, our group theoretically investigated the reaction mechanisms of the reversible photoswitching process in Dronpa and irreversible photoconversion in Kaede by the ONIOM(QM:MM) method, and proposed new mechanisms, i.e. photoisomerization coupled with ESPT in Donpa and E_{1cb} in Kaede.[15] Intrigued by the still unclear photo-decarboxylation process, one important advantage of PA-GFP is amazing contrast enhancement (~100-fold),9 compared to wt-GFP. The mechanistic picture will give helpful insight for understanding the experimental observation.

In this study, we investigated the reaction mechanism of the irreversible photo-induced decarboxylation of PA-GFP by the ONIOM(CASSCF:Amber) and ONIOM(CASPT2:Amber) methods.[16] Here, we focus on the most interesting and important process, i.e. the decarboxylation intrigued by the earlier charge transfer, while the last step leading to the anionic chromophore with the methyl capped decarboxylated Glu residue is beyond the scope of the present method and could be affected by the environment (solution). We will show that, upon UV or visible light irradiation, the neutral form of the chromophore is mainly excited to the first singlet excited state, an intra-molecular charge-transfer excited state (IntraCT) within the chromophore. As the C-C bond between CO_2^- and the remainder of GLU219 is elongated, the IntraCT state has to climb a hill and then reaches a transition state to an inter-molecular charge-transfer excited state (InterCT) originating from the crossing between the two diabatic states. Once in the InterCT state, the GLU219 C-C bond breaking is downhill in energy and the decarboxylation process takes place without barrier. The InterCT state continues to decrease its energy and then crosses with the so-called closed-shell state to produce CO_2 and the decomposition product in the ground state. The process leading to the

transition from the IntraCT to InterCT state is the rate-determining step of the overall decarboxylation reaction.

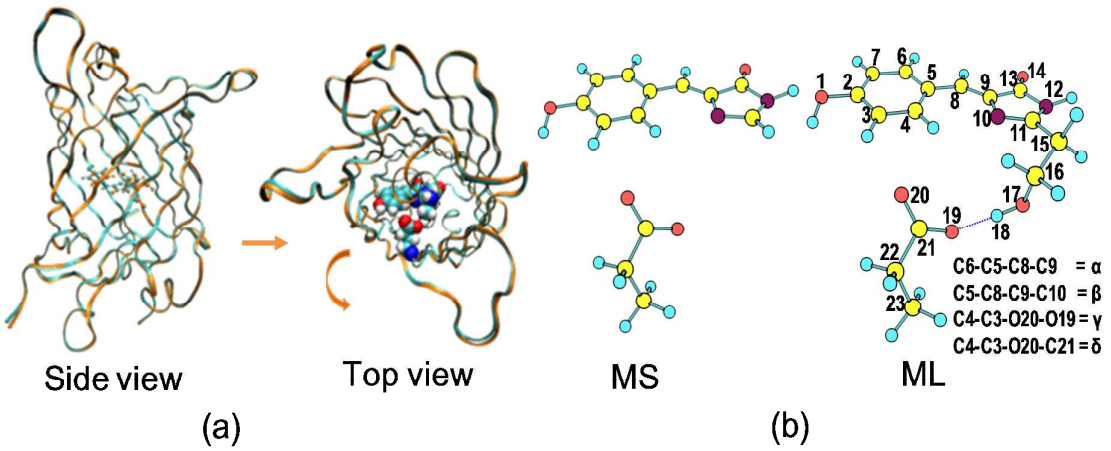
Computational Details

A. System Preparation and Classical Simulations

The initial PA-GFP structure with the neutral form of chromophore was obtained from Protein Data Bank (PDB ID: 3GJ1, 1.80 Å resolution).9 Residues in the chain A, the chromophore, and the corresponding 158 crystal water molecules were used. Hydrogen atoms were added to the protein by PDB2PQR program,[17] assuming the protonation state of all titratable residues at the experimental pH 8.0 on the basis of the results of PROPKA[18] program. The protonation state of histidine residues determined by PROPKA[18] analysis and visual inspection are: HID25, HIE77, HIE81, HID139, HID148, HID169, HID181, HID199, HID203, HID217, where HID and HIE stand for neutral histidines with protonated $N_{\delta 1}$ and $N_{\varepsilon 2}$ atoms, respectively. Meanwhile, CYS70 and CYM48 were also adopted, where CYS and CYM stand for the neutral and deprotonated states of a cysteine, respectively. Standard protonation state was used for the other amino acids. The missing atoms of the following residues (ARG80; ASP173; GLU132, 124, 90; LYS3, 26, 52, 79, 101, 107, 156, 158, 214; THR230) were mended using Amber9 program.[19] The overall charge of the system after adding missing atoms (in total 4049 atoms) is -7. The charged protein with the crystallographic water molecules was neutralized with sodium ions and fully solvated in a truncated octahedron water box of about 8 Å.

In order to obtain a good initial protein structure for ONIOM calculations, a series of constrained classical molecular mechanics (MM) optimizations were performed starting from the above-prepared protein structure with AMBER94 all-atom force field and TIP3P water model.[20] All available force field parameters, charges, and atom types of each amino acid have been taken from the AMBER library. RESP atomic charges of the chromophore (CRO) were derived at the HF/6-31G* level by Gaussian 09[21] and Amber9 program.[19] Additionally, similar Amber atom types derived from the Amber ANTECHAMBER module[22] were applied to the chromophore (Table S1 in Supporting Information (SI)). The hydrogen atoms, solvation waters and counterions, backbone atoms, all atoms in the MM region were

sequentially released in the optimizations, while active-site atoms in CRO, atoms of GLU219 involved in hydrogen-bond network around the active site such as SER202, WAT244 254 374 were kept fixed throughout the MM optimization. The non-bonded cutoff radius was set to 12 Å. In fact, superimposition of the X-ray and Amber-optimized geometries shown in Scheme 3 indicates an excellent agreement of these geometries.



Scheme 3. Superimposition of the X-ray (Cyan) and Amber optimized (Orange) geometries by VMD program[23]: (a) All heavy atoms (CPK representation) and (b) QM models (MS and ML with atomic numbering), with the atomic numbering system (viewed by Chemcraft package[24]).

In order to reduce the computational cost in the following ONIOM calculations, the solvation shell and counterions introduced by the AMBER Leap module were excluded from the MM-optimized structure. Finally, the entire system consists of 4046 atoms, including the protein, chromophore and 157 crystal water molecules, since one crystal water molecule was replaced by Na^+ by the AMBER Leap module (see Scheme 3a).

B. Gas Phase Model and ONIOM Calculations.

All QM (except for the CASSCF and CASPT2 methods) and ONIOM calculations were carried out with the Gaussian 09 program.[21] To describe different electronic states of the chromophore in the gas phase and protein, the complete-active-space self-consistent-field (CASSCF) and CASPT2 methods were also used, and energies and gradients were computed

externally by the Molcas74 program.[25]

B1. Gas phase model test calculations. Inspired by previous studies,[15,26] two different gas-phase models, MS (32 atoms) without a CH_2CH_2OH group in the chromophore and thus without hydrogen bond between CRO and GLU219, and ML (39 atoms) with CH_2CH_2OH and thus hydrogen bond (Scheme 3b), were set up. The QM model including more H-bonds may help to give more accurate results. However, the energy of IntraCT, the reactive state, is almost not affected by the H-bonding. The IntraCT in ML is S_1 which is in agreement with other calculation using much larger active space[15a,b] or larger model.[13c] Therefore, with or without the other H-bonding, the qualitative mechanistic picture should not be changed. Preliminary TD-CAM-B3LYP/6-31G*//CAM-B3LYP/6-31G*[27] calculation estimates that both models give similar relative excitation energies for IntraCT state, but the hydrogen bonding obviously makes the InterCT state more blue-shifted (Table S2, S3 and Figure S1). These calculations suggest that the ML model is more reasonable and that the IntraCT state should be the first electronic excited state.[15a,b] In addition, this hydrogen bonding included in ML model should also reduce basicity of GLU219. Here, four atoms (O1, N12, C22 and C23) are fixed during the optimization of the gas phase model.

The state-averaged SA6-CASSCF[28a,b] method with equal weights of six lowest-lying singlet states (S_0 to S_5)[28c] and 3-21G as well as 6-31G* basis sets were used for geometry optimization. Since we always use SA6-CASSCF, we simply call this CASSCF hereafter. As we know, the complete active space self-consistent field (CASSCF) wave function has sufficient flexibility to model the changes in electronic structure upon electronic excitation. However, the most important part of this application is to choose the proper active space based on the balance of the accuracy and high computational costs. Inspired by successful Martinez's small active space (2e/2o) for the GFP chromophore,[14h] an active space CAS(10e/7o) with 10 electrons in 7 orbitals was chosen to optimize the geometries, including one pair of π and π* orbitals (2e/2o) of the chromophore and n, π, π* orbitals (8e/5o) of the CO_2^- of GLU219 (see Figure S2). The exclusion of all the other π and π* orbitals of the chromophore is reasonable because these additional π→π* excitations were found to be in lower energy than the IntraCT state at the CASSCF level but not relevant to the present decarboxylation process (similar to our previous study[15a]). For instance, the

IntraCT state with the largest *f* value was found to be of S_4 at this CASSCF(10e/7o) level, which is found to be S_7 or even higher with larger active spaces, making the geometry-optimization unnecessarily expensive and complicated. Basically, the static correlation can be increased as much as possible by using the larger active space. But in present work, the reactive charge-transfer states (IntraCT) can be significantly affected by dynamic correlation that can be considered by second order perturbation theory method CASPT2. In present work, the IntraCT state is the first singlet excited state, S_1 (Table S4) after the correction by CASPT2 single-point energy calculations using the reference functions obtained by the above CASSCF method.[25] Here, an imaginary level shift of 0.1 au was applied to avoid intruder state problems,[29] since the latest IPEA Hamiltonian[30] was found to give a larger error in the vertical absorption energy than the original Hamiltonian.[15a] Also, both MS-CASPT2 and CASPT2 calculations gave similar wave functions and energy trend. Therefore, we mainly discuss CASPT2 calculations for simplicity. Additionally, to prove the insignificant contributions of the other orbitals, a larger active space for CASSCF/6-31G* and CASPT2/6-31G* single-point calculations were performed, such as the contribution from σ and σ* orbitals of the cleavage C-C bond by (12e/9o), or the contribution from one or two more pairs of π and π* orbitals of the chromophore by (12e/9o) or (14e/11o).

In a preliminary calculation using the TD-CAM-B3LYP method, the CO_2 group of GLU219 dissociates spontaneously on the InterCT excited state. Therefore, the bond between CO_2 and the rest of the GLU219 (C21-C22 bond in Scheme 3) was chosen as the key reaction coordinate. Approximate minimum energy path (MEP) for the decarboxylation involving the C21-C22 bond cleavage in S_1 in the gas phase was determined by relaxed scan optimizations with various C21-C22 distances. Since the order of states by the small basis set 3-21G is same with that by 6-31G* (Table S4), the MEP was determined at the CASSCF/3-21G level for S_1 with the ML model and the single-point CASPT2/3-21G//CASSCF energies were then plotted (Figure S3a). Here, two crossings were found. One involves a S_2/S_1 (IntraCT and InterCT) crossing and the other is the S_1/S_0 crossing; both should play important role in the photo-decarboxylation process. As shown in Figure S3b, we found a rotation along the single C-C bridge bond of the chromophore in

S_{1min} and other points near S_{1min} during the geometry optimization. Only when the C-C bond is significantly elongated along the InterCT state, two rings of the chromophore become roughly co-planar. These preliminary calculations in gas phase were used as the starting point for further studies in protein given below.

B2. ONIOM(QM:MM) calculations. In order to take into account the effect of protein environment on the QM calculations, the MM-optimized geometry by the classical simulations obtained in Section A was then refined by the following ONIOM calculations. The ONIOM calculations have been performed with a two-layer ONIOM(QM:MM) scheme,[16,31] in which the interface between the QM and MM regions is treated by hydrogen link atoms.[32] CASSCF as well as CAM-B3LYP methods, as mentioned above, and AMBER force field[20] are employed as the high and low level methods, respectively. It was previously found that an updated charge (UC) scheme combined with mechanical embedding (ME) scheme works well in describing geometries during the reaction, and we followed a six-step procedure[16i] using Mulliken charge, referred as the ONIOM(CASSCF:Amber)-MEUC method[16j] hereafter, for geometry optimization. The convergence criterion is that the total ONIOM energy difference between the last two rounds of the optimization is less than 0.1 kcal/mol. In optimization, the MM part was further divided into the optimized and the frozen regions.[33] The residues within ~8 Å of the chromophore were assigned to the optimized MM region and allowed to be optimized, while the others are fixed at the MM structure. Single-point calculations including dynamic correlation were performed at the ONIOM(CASPT2:Amber) level in an electronic embedding (EE) scheme at the ONIOM(CASSCF:Amber)-MEUC optimized geometries.

Relaxed scan optimization, as described in Section B1, was also employed to obtain the approximate minimum energy path (MEP) in the protein by ONIOM(CASSCF:Amber)-MEUC method. The MEP starts from the minimum of the S_1 state, S_{1min}, and continues with various C21-C22 distances on S_1. Single-point energies with dynamic correlation for all six states were calculated using the ONIOM(CASPT2:Amber)-EE method. Here we should note that different states have different charge distributions and polarizations. As shown later, a linearly interpolated path was used to describe the reaction process in the multi-dimensional space at a certain stage of

the reaction. Along the MEP, not only the C21-C22 bond distance but also some other coordinates change and participate in the reaction process.

Results and Discussion

A. Structures, Absorption and Emission.

The ground-state PA-GFP was optimized at the ONIOM(CASSCF:Amber)-MEUC level and this Frank-Condon (FC) structure is called S_{0min}. Major bonding parameters are shown in Figure 1 (see also Scheme 3). The computed geometry of the chromophore for this S_{0min} is similar to that for Dronpa which was optimized by the ONIOM(SA4-CASSCF(12e/12o)/ 6-31G*:Amber) method.[15a] The S_1 state is an intra-molecular charge-transfer (IntraCT) state within the chromophore, and the optimized structure, denoted as S_{1min}, shows some striking differences from S_{0min} in the geometrical parameters of the chromophore part. In particular, the bond-order inversion of the bridging C5-C8 and C8-C9 bonds was observed;[15a] the C5-C8 (1.45 Å) and C8-C9 (1.34 Å) bonds show considerable C-C and C=C bond character in S_{0min}, respectively,[15a] while the C5-C8 (1.36 Å) and C8-C9 (1.46 Å) bonds represent C=C and C-C bonds in S_{1min}, respectively.[15a]

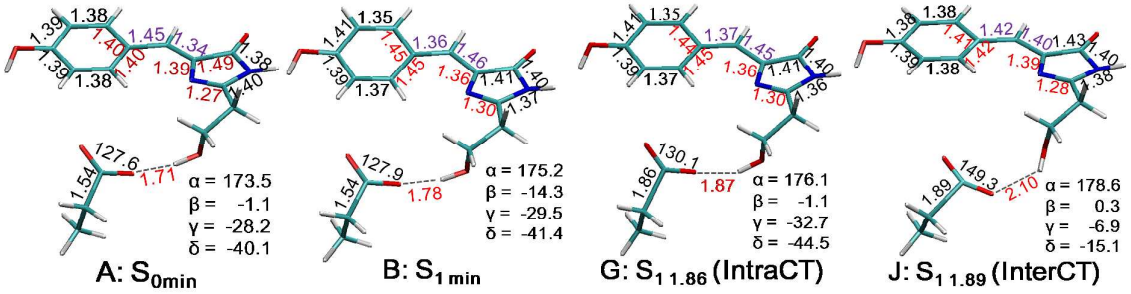


Figure 1: Key geometrical parameters (in Å and deg) of the minima on S_0 (A: S_{0min}) and S_1 (B: S_{1min}) states, Point G (~ $\mathbf{X_{IntraCT/InterCT}}$) and Point J (~ $\mathbf{X_{InterCT/CS}}$) optimized by the ONIOM(SA6-CASSCF(10e/7o)/6-31G*:Amber)-MEUC method.

Formally, the IntraCT state represents an intra-molecular electron transfer from the phenolic ring to the imidazolidinone ring of the chromophore,[15a] and the InterCT state represents an inter-molecular electron transfer from GLU219 to the chromophore. Mulliken charge distribution analysis was shown in Figure 2 (also Figure S6), where the QM part of the system was divided into three parts. Comparing with the ground state, the excitation to

the IntraCT state leads to about 0.21 e transferred from the phenolic ring (Part 1) to the imidazolidinone ring (Part 2) of the chromophore, while keeping GLU219 negatively charged. On the other hand, the excitation to the InterCT state induces about 0.96 e transferred from GLU219 (Part 3) to the chromophore (more to the imidazolidinone ring), leading GLU219 to be nearly neutral. Accordingly, the calculated dipolar moment of the chromophore is significantly increased by ~4.5 Debye from S_0 to the InterCT state, supporting that dipole change is accompanied with the charge-transfer excited state.[12b]

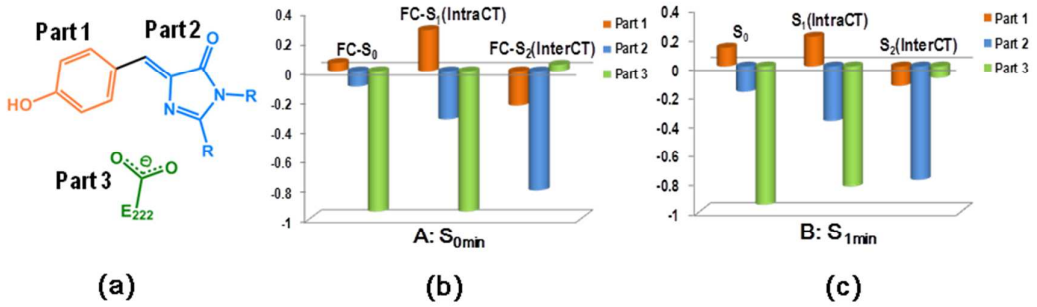


Figure 2: Mulliken charge distribution analysis for three parts (defined in (a)) by the ONIOM(SA6-CASSCF(10e,7o)/6-31G*:Amber)-MEUC method for S_0, S_1 (IntraCT) and S_2 (InterCT) states at (b) S_{0min} and (c) S_{1min} structures.

The vertical absorption energies at the S_{0min} structure and emission energies at the S_{1min} structure for PA-GFP at the ONIOM(CASPT2:Amber) level are collected in Table 1, in which zero point energy corrections (ZPC) were not included. The vertical absorption energy to S_1 (i.e. IntraCT state) is about 79.4 kcal/mol by the ONIOM(CASPT2:Amber)-EE method.9 The vertical emission energy from S_{1min} is roughly 64.8 kcal/mol by the ONIOM(CASPT2:Amber)-EE method (see Table S5 for the other results). These values without ZPC are a little higher than the experimental values for PA-GFP (72.1 and 55.6 kcal/mol for absorption and emission, respectively),9 but they are very similar to our previous work (83.5 and 68.7 kcal/mol) estimated by much larger active space for Dronpa.[15a] The adiabatic absorption energy from S_{0min} to S_{1min} is calculated to be around 73.4 kcal/mol by the ONIOM(CASPT2:Amber)-EE method. The energy difference between S_2 (InterCT) and S_1 (IntraCT) states is about 32 kcal/mol at both S_{0min} and S_{1min} structures, suggesting that

their potential surfaces in this geometrical region are nearly parallel to each other. The overestimation about the vertical absorption energies and emission energies at present approaches (such as method and QM model) can be due to shortcomings of imperfect QM model or incomplete active space in CASSCF or different protonated residues in MM region.[13f,g] Although a computationally demanding approach such as larger QM model or larger active space in CASSCF is highly desirable for obtaining accurate results,[13c] it could not change the mechanistic picture of the photo-decarboxylation, which is the primary goal of this study.

Table 1. Vertical absorption energies (with oscillator strength, *f*) at S_{0min} and relative energies (relative to S_0 at S_{0min})[a] at S_{1min} (kcal/mol) by the ONIOM(SS-CAS(10e/7o)PT2/6-31G*:Amber)-EE at the ONIOM(SA6-CASSCF(10e/7o) /6-31G*:Amber)-MEUC-optimized S_{0min} and S_{1min} geometries

		Vertical absorption energies at S_{0min} (FC)		Relative energy at S_{1min}
State	Nature[b]	E	*f*	E
S_4	InterCT" p	116.1	0.00011	105.5
S_3	InterCT' m-n	113.6	0.00036	103.7
S_2	InterCT	111.3	0.0013	105.4
S_1	IntraCT	79.4	1.28	73.4
S_0	Ground state	0.0		8.6

[a] Without ZPC. [b] "Qualitative description of nature" of states is defined in Table S2.

B. Photo-induced Decarboxylation Reaction.

The approximate MEP for the decarboxylation process on the S_1 state (IntraCT state) starting from S_{1min} was determined by a relaxed scan with various C21-C22 bond distances (i.e. decarboxylation coordinate, $R_{C21-C22}$) from 1.54 Å to 1.88 Å at the ONIOM(CAS:MM)-MEUC level. One of the optimized structures (point G) is given in Figure 1 (the rest of points B, C, D, E, F, G, H and I are given in Figure S4). The relative energies of S_0, S_1 and S_2 states were refined by the ONIOM(CASPT2:Amber)-EE method, as shown in Figure 3A (see also Figure S5-S7 and Table S6-S7). As the C21-C22 distance increases, the energy of S_1 (IntraCT, blue lines) increases gradually, while that of S_2 (InterCT, red lines) decreases in both the CASSCF and CASPT2 methods. The IntraCT excited state forms a conical intersection to connect to the InterCT excited state, when the C-C distance

increases to be a little beyond point F during the first stage of the decarboxylation process; the ONIOM(CAS:MM)-MEUC energy of the InterCT state was higher than that of the IntraCT state by 0.4 kcal/mol at point F ($R_{C21-C22}$ = 1.85 Å), which then became to be lower in energy by about -1.0 kcal/mol at point G ($R_{C21-C22}$ = 1.86 Å). The crossing of the two states (S_1 and S_2) generates an adiabatic transition state on the lower surface. In this work, we did not determine the exact location of this adiabatic transition state or crossing, called $X_{IntraCT/InterCT}$, but we took point G as an approximate transition state $X_{IntraCT/InterCT}$ for the further C21-C22 dissociation on the S_1 adiabatic state. Inclusion of dynamic correlation lowers the energies of both S_1 and S_2 relative to S_0. The IntraCT-InterCT energy difference at point G by the ONIOM(CASPT2:MM)-EE method was around -2.5 kcal/mol, still small enough for qualitative discussion of the crossing. The relative energy of IntraCT at point G (approximate $X_{IntraCT/InterCT}$) is 13.4 and 12.3 kcal/mol above that of the same state at the FC point by the ONIOM(CAS:MM)-MEUC and ONIOM(CASPT2:MM)-EE methods, respectively. This energy increase constitutes a barrier for the decarboxylation on the S_1 state and turns out to be the rate-determining barrier for the entire decarboxylation process.

Along the S_1 pathway from point B (S_{1min}) to point G ($\sim X_{IntraCT/InterCT}$) as the C21-C22 bond is gradually elongated from 1.54 to 1.85 Å, the O20-C21-O19 bond angle of GLU219 slightly increases from 127.9° to 130.1° and the H-bond distance (H18-O19) increases from 1.78 Å to 1.88 Å. The dihedral angles α and β increase from 175.2° to 176.1° and from -14.3° to -1.4°, respectively, indicating that the chromophore part becomes nearly coplanar. Meanwhile, the other geometric parameters did not change much from those of S_{1min}.

After the transition *via* point G ($\sim X_{IntraCT/InterCT}$), the reaction proceeded along InterCT state, as shown in Figure 3(B and C). As expected, the optimized structures and energies obtained on the InterCT state were largely different from these on the IntraCT state due to the nature change of the states (Table S6 and Figure S4). The structure of point J is given in Figure 1.

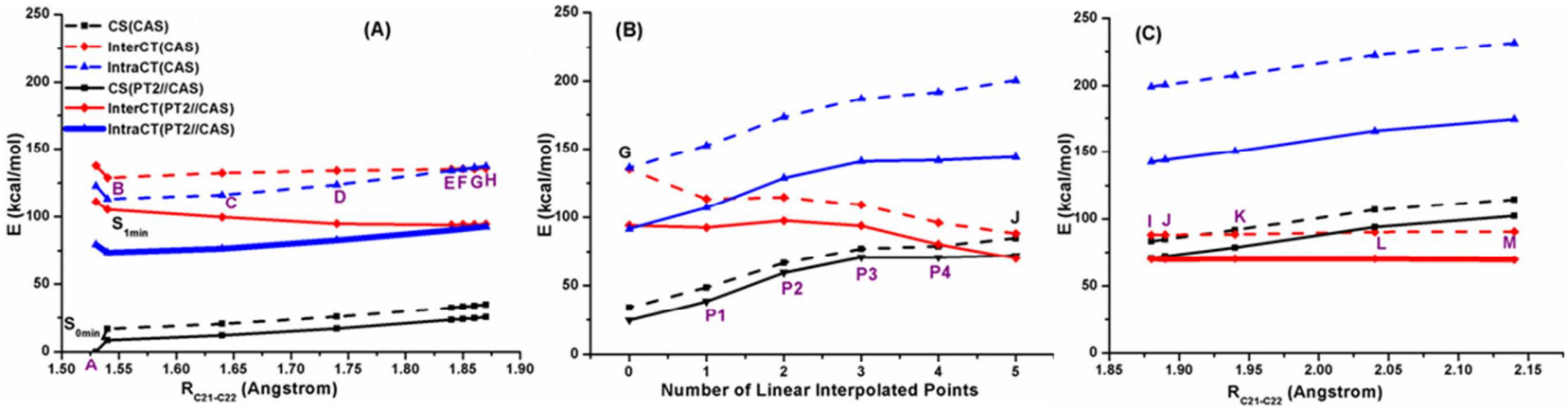


Figure 3. PES for the C-C bond cleavage (A) from S_{1min} to point H (near the transition) along the IntraCT state as S_1; (B) A linearly interpolated pathway between point G and point J was shown; (C) from point I to point M along InterCT state as S_1 optimized by the ONIOM(SA6-CASSCF(10e,7o)/6-31G*:Amber)-MEUC method. Single-point energies were refined by the ONIOM(CASPT2(10e,7o)PT2/6-31G*: Amber)-EE method.

There are substantial differences in both geometries and energies between point G (on the IntraCT MEP with $R_{C21\text{-}C22}$ = 1.86 Å) and point J (on the InterCT MEP with $R_{C21\text{-}C22}$ = 1.89 Å). Going from point G to point J, as the ONIOM(CASPT2:MM)-EE level shown in Figure 3B (and Table S6), the relative energies of the IntraCT and S_0 states increase very significantly by 52 and 46 kcal/mol, respectively, while that of the InterCT state drops by 24 kcal/mol. These are due to large geometrical differences between the IntraCT MEP and InterCT MEP. As seen in Figure 1, the bridge C5-C8 bond is increased by 0.05 Å and the C8-C9 bond is decreased by 0.05 Å from point G to point J. Qualitatively, the HOMO and LUMO of the chromophore in the IntraCT state contain one electron each, while in the InterCT state HOMO contains two electrons and LUMO one electron. The bond distance changes for G $\rightarrow$ J above reflect the C5-C8 anti-bonding and C8-C9 bonding nature of the chromophore HOMO. More importantly the carboxyl bond angle O20-C21-O19 increased from 130.1° to 149.3°, reflecting that while CO_2 is formally an anion in the IntraCT state, the neutral CO_2 prefers to be collinear in the InterCT state. At the same time, the H-bonding becomes longer and the dihedral angles γ and δ become much closer to zero, indicating that the chromophore and the OCO group become more coplanar. The Mulliken population analysis (Figure S6) is consistent with the present picture.

Due to the existence of tens kcal/mol difference near 1.85 Å, the PES of this region is considered by a linearly-interpolated pathway (Figure 3(B)). The reason is that near the crossing point, more coordinates other than one C21-C22 are involved in the intermolecular electron transfer. Here, the linearly-interpolated approach was employed in order to trace the potential energy surface from point G (IntraCT state) to point J (InterCT state). Single-point energies of the InterCT, IntraCT and CS states at these interpolated points as well as points G and J are depicted in Figure 3B (as well as Table S6b). The InterCT potential energy surface from point G seems to go over an insignificant barrier of only about 3.2 kcal/mol at point P2, which are 97.5 and 94.2 kcal/mol for point P2 and point G by the ONIOM(PT2:MM)-EE method, respectively, relative to S_{0min}. After this point the InterCT energy decreases rather quickly to reach point J (InterCT state).

As the continuous pathway on the InterCT state (shown in Figure 3C), the relaxed scan at the ONIOM(CAS:MM)-MEUC level was done along C21-C22 from 1.88 (point I) to 2.14

Å (point M). As depicted in Figure 3C, the energy of what we called the closed-shell (CS) state increased steadily and substantially, while that of the InterCT state is nearly unchanged. This demonstrates the fact that at the limit of C21-C22 bond broken, the InterCT state represents the chromophore anion plus CO_2, while what we called the CS state corresponds to the chromophore plus CO_2^-. The dissociation of CO_2 on the InterCT state seems to be barrierless and is slightly endothermic. Along the relaxed scan from point I to point M, the main geometrical change occurred gradually in the angle O20-C21-O19 from 148.7° to 161.3°. Especially for the last two points, large changes in γ and δ indicate the loss of H-bonding to CO_2, which becomes more freely movable.

At point J with $R_{C21-C22}$ of 1.89 Å, the energy of InterCT state is higher than that of the CS state by 3.4 kcal/mol, but at the point K with $R_{C21-C22}$ of 1.90 Å the energy of the InterCT state is *lower* than that of the CS state by -3.3 kcal/mol at ONIOM(CAS:MM)-MEUC level. In this vicinity the two states cross via a crossing $X_{InterCT/CS}$. Again, we did not determine the exact structure of this CI, but we considered point J as an approximate $X_{InterCT/CS}$. Obviously, the dynamically correlated ONIOM(CASPT2:MM)-EE method lowers the energies of both states relative to the reference S_{0min}; at point J the energy of InterCT is lower by -2.1 kcal/mol than that of the CS state, which is still acceptable qualitatively as the approximate $X_{InterCT/CS}$. The Mulliken population analysis (Figure S6) is consistent with the present picture.

Notably, similar trends (see Figure S8, Figure S11-S12 as well as Table S8-10) obtained by single-point CASSCF/6-31G* and CASPT2/6-31G* calculations with larger active space supports the same mechanistic picture of the photo-decarboxylation by the small active space (10e/7o), which proved only a small contribution from σ and σ* orbitals of the cleavage C-C bond or from one or two more pairs of π and π* orbitals of the chromophore.

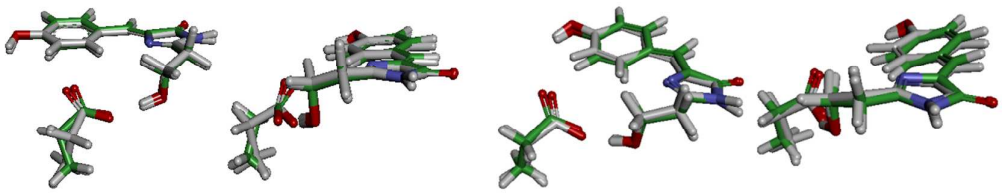
C. Effects of the Protein Environment on the Decarboxylation Reaction.

As shown in Figure S4, the geometries of the Ph-ring and N-ring in all chromophores optimized in the protein are roughly coplanar. However, rotation along the single bridge bond (C8-C9) occurred in the IntraCT state in our gas-phase model test calculations (see the computational details and Figure S3a and b). We have also re-optimized the geometries of

the QM part at point A, B, J and M by removing the MM protein part.

The superimposition of the QM model between the gas-phase and ONIOM results for S_{0min} and S_{1min} are given in Figure 4. The carboxylic group rotates in the protein, compared to that in gas-phase model calculations. Flexibility of the carboxylic group is in accordance with the experimental proposal that the orientation of the carboxylic group can be changed by mutating Thr to His leading to the enhanced emission in PA-GFP. Except that, the gas-phase ML model can overlay well with the QM model of the ONIOM and X-ray structures (see Figure S10 in SI). Such qualitatively good agreement suggests that the gas-phase ML model in S_0 is a reasonable starting point to study PA-GFP.

Unfortunately, the Ph-ring along the single bridge bond (C8-C9) easily rotates in the gas-phase model for S_{1min} (β: -30.6/-14.3, see Figure 4B), presumably due to no or a very low barrier of the photosiomerization.[14e-f,15a,c] That made MCSCF convergence difficult, due to a small gap between S_0 and S_1, implying a probable branching route to S_0 rather than the decarboxylation. In this connection, recent non-adiabatic QM/MM molecular dynamics simulations showed an ultra-fast radiationless decay for the related neutral chromophore of the other fluorescence proteins.[14e,15c] Therefore, one possible protein effect is to increase the barrier of the photoisomerization. On the other hand, deformation energy of the QM model within the protein at point A, B, J and M was found to be quite similar (~10-11 kcal/mol), indicating that the strain effect by the protein is not important in the decarboxylation reaction.



(A) S_{0min} (β: -1.1/-1.1) (B) S_{1min} (β: -30.6/-14.3)

Figure 4 Superposition between Gas ML by CASSCF/6-31G* method and QM model by ONIOM(CASSCF/6-31G*:Amber)-MEUC. (A) Front and side views of Gas ML (Green) and QM model at S_{0min}. (B) Front and side views of Gas ML (Green) and QM model at S_{1min}.

Table 2 Total Energies (au), Deformation Energies (kcal/mol)

Point	Deformation
A	10.8
B	10.5
J	11.8
M	9.7

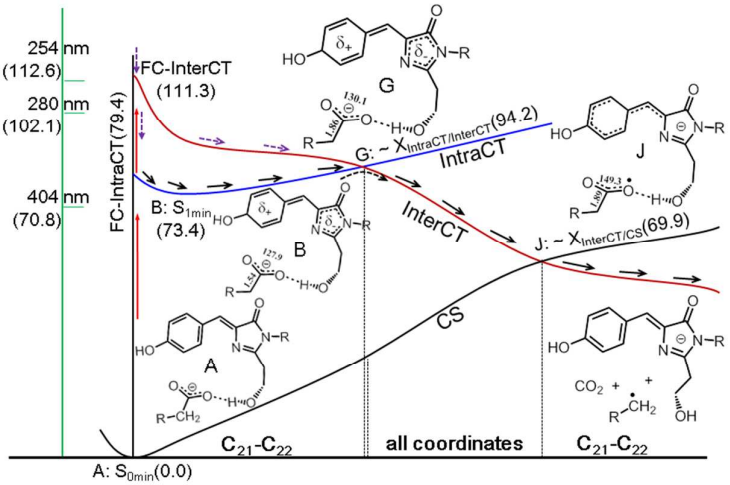
Energetic contribution from electrostatic (ES) and VDW interactions with different residues of the protein environment was also evaluated (Figures S13-14), where B-A and G-A correspond to the energy difference between IntraCT state and S_{0min}, and J-A and M-A correspond to the energy difference between InterCT state and S_{0min}. From Figure S13A and S13B, the electrostatic interaction difference analysis indicated that charged residues Gln91 and Arg93 mainly stabilize the IntraCT state, since the excitation makes the N-ring more negatively-charged. A much larger electrostatic stabilization to the InterCT excited state by Arg93 and Gln91 was also observed (Figure S13C and D), as the N-ring becomes even more negatively-charged in the InterCT than that in the IntraCT state. This is in accordance with the Mulliken charge analysis (Figure S6). In comparison, there is insignificant VDW stabilization to the IntraCT interactions relative to S_0. The main VDW stabilization to the InterCT interactions by Ser202 relative to S_0 was found to slightly increase to about 4-5 kcal/mol. Therefore, the overall effect of ES and VDW were found to stabilize the IntraCT and InterCT states more than S_{0min}.

D. Reaction Mechanism of the Photo-induced Decarboxylation.

On the basis of the present calculations (Figure 3), the photo-induced decarboxylation mechanism in PA-GFP can be summarized as follows. At a low energy excitation, such as 404 nm (i.e. 70.8 kcal mol^{-1}) irradiation, the chromophore is at first excited to S_1, an IntraCT state within the chromophore. The S_1 state (blue curves in Figure 3A) goes up the transition state or $X_{IntraCT/InterCT}$ in the vicinity of point G, where the S_1 state adiabatically but quickly changes its character from the IntraCT to InterCT state. The energy required to reach point G is about 94 kcal/mol at the more reliable ONIOM(CASPT2:MM)-EE level, relative to S_{0min}. This is the rate-determining barrier of the entire decarboxylation reaction. Therefore one-photon irradiation at low energy such as 404nm does not result in decarboxylation[13b,c,d], in agreement with the 2- or 3-photon decarboxylation process *via* Kolbe-type mechanism at

low energies observed in the experiments.[13e] For 254 and 280 nm radiation (i.e. 112.6 and 102.1 kcal mol^{-1}, respectively) also used experimentally, the excitation to the higher electronic states, like S_2, S_3 and S_4, all InterCT-type states, would also take place,[7b] as well as the non-FC excitation to S_1 with internal vibration excitation. Generally,[34] the excited chromophore in these higher electronic states may decay to IntraCT state (S_1) via ultrafast internal conversion, generating excess internal energy. These excess internal energies will allow the chromophore to reach point G on the IntraCT state and *adiabatically* switch to the InterCT state to continue decarboxylation. Some of the chromophores excited to S_2, S_3 or S_4 state might directly follow the repulsive InterCT surface without barrier (red curves in Figure 3A) to reach the vicinity of point G and continue *diabatically* onto the still repulsive InterCT state. In either way, the chromophore follows the InterCT state from $X_{IntraCT/InterCT}$ and the decarboxylation takes place (Scheme 4) on the InterCT state, which becomes the ground state after $X_{InterCT/CS}$. These findings are consistent with the experimental observation[7b] that the rate of the decarboxylation depends on the light resource and was found to be ultrafast in WT-GFP (an estimated activation barrier of ~0.2 kcal/mol only) using 254 nm radiation.[7a]

The charge-transfer nature of the IntraCT and InterCT states was characterized by Mulliken charge analysis (Figures 2 and S6). The electron deficiency in the phenolic ring after the intra-molecular charge transfer excitation makes it easier to accept an electron from the GLU anion which can easily decarboxylate in the neutral form. This is in agreement with the proposed Kolbe mechanism for the photo-induced decarboxylation.



Scheme 4. The putative schematic representation with the relative energies by ONIOM(CASPT2:Amber)-EE (kcal/mol) as well as key structural features for the photo-induced decarboxylation pathway

Conclusions

In the present study, we reported the first theoretical study on the photo-induced irreversible decarboxylation of PA-GFP by ONIOM(CASSCF:Amber) and ONIOM(CASPT2:Amber) methods. As the decarboxylation coordinate (C20-C21 distance) increases, the first singlet excited state (S_1, IntraCT) with the largest transition intensity was found to increase its energy and reach the approximate transition state $X_{IntraCT/InterCT}$ with a barrier of ~ 12.3 kcal/mol. Although this value may depend on the method used, the qualitative description of the mechanistic picture should be unchanged. The rate-determining transition state ($X_{IntraCT/InterCT}$) is the result of avoided crossing between the IntraCT and InterCT states and the reaction proceeds adiabatically on the S_1 state that now is the InterCT state. This provides an efficient decay channel to connect the bright state and the dark state. As the C20-C21 distance increases further, the InterCT state with virtually no barrier goes through the crossing $X_{InterCT/CS}$ diabatically and becomes the ground state. The process after $X_{IntraCT/InterCT}$ proceeds without barrier and gives the product along the Kolbe-type decarboxylation mechanism. The comparison of deformation geometries and energies shows that the strain effect by the protein is not important in the decarboxylation reaction, but may decrease the photoisomerization possibility. Among all the residues, two charged residues Gln91 and Arg93 mainly stabilize the IntraCT and InterCT states via the electrostatic interaction, while Ser202 slightly stabilizes the InterCT by the VDW interaction relative to S_{0min}. Therefore, the overall effect of ES and VDW were found to stabilize the IntraCT and InterCT states more than S_{0min}. From the present study, ultrafast decarboxylation *via* one higher-energy photo-irradiation or multiple lower-energy photo-irradiation involving bright states and inter-molecular charge-transfer state could be a common phenomenon for different photoactivatable fluorescent proteins having similar active sites.[35]

Acknowledgment. We thank Dr. Fengyi Liu for discussions and assistance on excited states

and MOLCAS calculations, Dr. Xin Li for fluorescent proteins discussions and bringing PA-GFP to our attention, and Dr. Miho Hatanaka for her code of energy decomposition calculations. L.W.C. acknowledges the FIFC Fellowship. This work is supported in part 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. The computational resources at Research Center of Computer Science (RCCS) of the Institute for Molecular Science (IMS) and at ACCMS of Kyoto University are acknowledged.

Supporting Information Available: Complete refs. 19, 21 and 25, Figures S1–S12, Tables S1-S10, and Cartesian coordinates of the QM-optimized structures. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) (a) Tsien, R. Y. *Annu. Rev. Biochemistry* **1998**, *67*, 509-544. (b) Zimmer, M. *Chem. Rev.* **2002**, *102*, 759-782. (c) Shaner, N. C.; Patterson, G. H.; Davidson, M. W. *J. Cell Sci.* **2007**, *120*, 4247–4260.

(2) (a) Patterson, G. H.; Lippincott-Schwartz, J. *Science* **2002**, *297*, 1873-1877. (b) Chudakov, D. M.; Verkhusha, V. V.; Staroverov, D. B.; Souslova, E. A.; Lukyanov. S.; Lukyanov, K. A. *Nat. Biotechnol*. **2004**, *22*, 1435–1439. (c) Subach, F. V.; Patterson, G. H.; Renz, M.; Lippincott-Schwartz, J.; Verkhusha, V. V. *J. Am. Chem. Soc.* **2010**, *132*, 6481-6491. (d) 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*, 191-194. (d) Verkhusha, V. V.; Sorkin, A. *Chem. Biol.* **2005**, *12*, 279–285. (e) Subach, F. V.; Malashkevich, V. N.; Zencheck, W. D.; Xiao, H.; Filonov, G. S.; Almo, S. C.; Verkhusha, V. V. *Proc. Natl. Acad. Sci U.S.A.* **2009**, *106*, 21097-210102. (f) Subach, F. V.; Patterson, G. H.; Manley, S.; Gillette, J. M.; Lippincott-Schwartz, J.; Verkhusha, V. V. *Nat. Methods.* **2009**, *6*, 153-159.

(3) (a) Kaede: Ando, R.; Hama, H.; Yamamoto-Hino, M.; Mizuno, H.; Miyawaki, A. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 12651–12656. (b) Mizuno, H.; Mal, T. K.; Tong, K. I.; Ando, R.; Furuta, T.; Ikura, M.; Miyawaki, A. *Mol. Cell.* **2003**, *12*, 1051–1058. (c) Hayashi, I.; Mizuno, H.; Tong, K. I.; Furuta, T.; Tanaka, F.; Yoshimura, M.; Miyawaki, A.; Ikura, M. *J. Mol. Biol.* **2007**, *372*, 918–926. EosFP: (d) 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*, 15905–15910. (e) Nienhaus, G. U.; Wiedenmann, J.; Nar, H. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 9156–9159. (f) KikGR: Tsutsui, H.; Shimizu, H.; Mizuno, H.; Nukina, N.; Kuruta, T.; Miyawaki, A. *Chem. Biol.* **2009**, *16*, 1140–1147. (g) IrisFP: 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*, 18343–18348. (h) Fuchs, J.; Böhme, S.; Oswald, F.; Hedde, P. N.; Krause. M.; Wiedenmann, J.; Nienhaus, G. U. *Nat Methods*. **2010**, *7*, 627-630. (i) Subach, O. M.; Patterson, G. H.; Ting, L. M.;

Wang, Y.; Condeelis, J. S.; Verkhusha, V. V. *Nat Methods*. **2011**, *8*, 771-777. Dendra2 (j) 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*, 461–465.

(4) Dronpa: (a) Ando, R.; Mizuno, H.; Miyawaki, A. *Science* **2004**, *306*, 1370–1373. (b) Habuchi, S.; Ando, R.; Dedecker, P.; Verheijen, W.; Mizuno, H.; Miyawaki, A.; Hofkens, J. *Proc. Natl. Acad. Sci. U. S. A*. **2005**, *102*, 9511-9516. (c) Fron, E.; Flors, C.; Schweitzer, G.; Habuchi, S.; Ando, R.; De Schryver, F. C.; Miyawaki, A.; Hofkens, J. *J. Am. Chem. Soc.* **2007**, *129*, 4870–4871. (d) Mizuno, H.; Mal, T. K.; Wälchli, M.; Kikuchi, A.; Fukano, T.; Ando, R.; Jeyakanthan, J.; Taka, J.; Shiro, Y.; Ikura, M.; et al. *Proc. Natl. Acad. Sci. U. S. A*. **2008**, *105*, 9927–9932. (e) 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*, 13005–13009. (f) Andresen, M.; Stiel, A. C.; Fölling, J.; Wenzel, D.; Schönle, A.; Egner, A.; Eggeling, C.; Hell, S.W.; Jakobs, S. *Nat. Biotechnol*. **2008**, *26*, 1035–1040. (g) Brakemann, T.; Weber, G.; Andresen, M.; Groenhof, G.; Stiel, A. C.; Trowitzsch, S.; Eggeling, C.; Grubmüller, H.; Hell, S. W.; Wahl, M. C.; Jakobs, S. *J. Biol. Chem*. **2010**, *285*, 14603-14609. (h) Regis Faro, A.; Carpentier, P.; Jonasson, G.; Pompidor, G.; Arcizet, D.; Demachy, I.; Bourgeois, D. *J. Am. Chem. Soc*. **2011**, *133*, 16362-16365. (i) asFP595: 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*, 13070–13074. (j) mTFP0.7: Henderson, N. J.; Ai, H.-W.; Campbell, R. E.; Remington, S. J. *Proc. Natl. Acad. Sci. U. S. A*. **2007**, *104*, 6672–6677. (k) Bizzarri, R.; Serresi, M.; Cardarelli, F.; Abbruzzetti, S.; Campanini, B.; Viappiani, C.; Beltram, F. *J. Am. Chem. Soc*. **2010**, *132*, 85-95. (l) Grotjohann, T.; Testa, I.; Leutenegger, M.; Bock. H.; Urban, N. T.; Lavoie-Cardinal, F.; Willig, K. I.; Eggeling, C.; Jakobs, S.; Hell, S. W. *Nature* **2011**, *478*, 204-208. (m) Brakemann, T.; Stiel, A. C.; Weber, G.; Andresen, M.; Testa, I.; Grotjohann, T.; Leutenegger, M.; Plessmann, U.; Urlaub, H.; Eggeling, C.; Wahl, M. C.; Hell, S. W.; Jakobs, S. *Nat. Biotechnol.* **2011**, *29*, 942-947. (n) Chudakov, D. M.; Feofanov, A. V.; Mudrik, N. N.; Lukyanov, S.; Lukyanov, K. A. *J. Biol. Chem*. **2003**, *278*, 7215-7219.

(5) (a) Chalfie, M.; Tu, Y.; Euskirchen, G.; Ward, W. W.; Prasher, D. C. *Science* **1994**, *263*, 802-805. (b) Cubitt, A. B.; Heim, R.; Adams, S. R.; Boyd, A. E.; Gross, L. A.; Tsien, R. Y. *Trends. Biochem. Sci*. **1995**, *20*, 448-455.

(6) (a) Chattoraj, M.; King, B. A.; Bublitz, G. U.; Boxer, S. G. *Proc. Natl. Acad. Sci. U. S. A*. **1996**, *93*, 8362-8367. (b) Bublitz, G.; King, B. A.; Boxer, S. G. *J. Am. Chem. Soc.* **1998**, *120*, 9370-9371. (c) Creemers, T. M. H.; Lock, A. J.; Subramaniam, T. M.; Völker, S. *Nat. Struct. Mol. Biol*. **1999**, *6*, 557-560. (d) Brejc, K.; Sixma, T. K.; Kitts, P. A.; Kain, S. R.; Tsien, R. Y.; Ormö, M.; James Remington, S. *Proc. Natl. Acad. Sci. U. S. A*. **1997**, *94*, 2306-2311. (e) Palm, G. J.; Zdanov, A.; Gaitanaris, G. A.; Stauber, R.; Pavlakis, G. N.; Wlodawer, A. *Nat. Struct. Biol*. **1997**, *4*, 361-365. (f) van Thor, J. J. ; Zanetti, G.; Ronayne, K. L.; Towrie, M. *J. Phys. Chem. B*. **2005**, 109, 16099-16108. (g) Stoner-Ma, D.; Jaye, A. A.; Matousek, P.; Towrie, M.; Meech, S. R.; Tonge, P. J. *J. Am. Chem. Soc*. **2005**, *127*, 2864-2865. (h) van Thor, J. J.; Pierik, A. J.; Nugteren-Roodzant, I.; Xie, A.; Hellingwerf, K. J. *Biochemistry* **1998**, *37***,** 16915-16921.

(7) (a) van Thor, J. J.; Gensch, T.; Hellingwerf, K. J.; Johnson, L. N. *Nat. Struct. Mol. Biol.* **2002**, *9*, 37-41. (b) Bell, A. F.; Stoner-Ma, D.; Wachter, R. M.; Tonge, P. J. *J. Am. Chem. Soc*. **2003**, *125*, 6919-6926.

(8) (a) van Thor, J. J.; Georgiev, G. Y.; Towrie, M.; Sage, J. T. *J. Biol. Chem.* **2005**, *280*, 33652-33659. (b) Morise, H.; Shimomura, O.; Johnson, F. H. Winant, J. *Biochemistry* **1974**, *13*, 2656-2662. (c) Ward, W. W.; Cody, C. W.; Hart, R. C.; Cormier, M. J. *Photochem. Photobiol*. **1980**, *31*, 611-615.

(9) Henderson, J. N.; Gepshtein, R.; Heenan, J. R.; Kallio, K.; Huppert, D.; Remington, S. J. *J. Am. Chem. Soc*. **2009**, *131*, 4176-4177.

(10) (a) Kolbe, H. *Ann. Chim. Pharm.* **1849**, *69*, 257-294. (b) Vijh, A. K.; Conway, B. E. *Chem. Rev* **1967**, *67*, 623-664. (c) Martins, B. M.; Blaser, M.; Feliks, M.; Ullmann, G. M.; Buckel, W.; Selmer, T. *J. Am. Chem. Soc*. **2011**, *133*, 14666-14674.

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

(12) (a) Xu, Y.; Chen, X.; Fang, W.-H.; Phillips, D. L. *Org. Lett.* **2011**, *13*, 5472–5475. (b) Ding, L.; Chen, X.; Fang, W.-H. *Org. Lett.* **2009**, *11*, 1495-1498. (c) Ding, L.; Fang, W.-H. *J. Org. Chem*. **2010**, *75*, 1630-1636. (d) Ding, W. J.; Ni, L. Y.; Fang, W. H.; Yu, J. G. *J. Theor. Comput. Chem*. **2005**, *4*, 725-736.

(13) (a) Roy, A.; Field, M. J.; Adam, V.; Bourgeois, D. *J. Am. Chem. Soc.* **2011**, *133*, 18586–18589. (b) Hasegawa, J.Y.; Fujimoto, K.; Swerts, B.; Miyahara, T.; Nakatsuji, H. *J. Comput. Chem*. **2007**, *28*, 2443-2452. (c) Grigorenko, B. L.; Nemukhin, A. V.; Morozov, D. I.; Polyakov, I. V.; Bravaya, K. B.; Krylov, A. I. *J. Chem. Theory. Comput.* **2012**, 8, 1912-1920. (d) One of the referee suggested that the decarboxylation reaction could start from the anionic form induced by excited-state proton transfer, which is not excluded by this work. (e) Langhojer, F.; Dimler, F.; Jung, G.; Brixner, T. *Biophys. J*. **2009**, 96, 2763-2770. (f) Filippi, C.; Buda, F.; Guidoni, L.; Sinicropi, A. *J. Chem. Theory Comput*. **2012**, 8, 112–124. (g) One of the referee commented that the overestimated vertical and adiabatic S_1-S_0 energies may also be attributed to "MM deficiencies, like protein electrostatic response or protonation states that may be different from the ones postulated in this work or that may change during the reaction". We do not exclude a possibility that the protonation states can be changed during the entire decarboxylation reaction. However, excitation and relaxation to S_1 should be too fast to change the protonation states.

(14) Recent theoretical studies on reaction mechanism of GFP: (a) Bravaya K. B.; Subach O. M.; Korovina, N.; Verkhusha V. V.; Krylov A. I. *J. Am. Chem. Soc.* **2012**, *134*, 2807-2814. (b) Cui, G.; Lan, Z.; Thiel, W. *J. Am. Chem Soc.* **2012**, *134*, 1662-1672. (c) Sanchez-Garcia, E.; Doerr, M.; Hsiao, Y. W.; Thiel, W. *J. Phys. Chem. B* **2009**, *113*, 16622-31. (d) Filippi, C.; Buda, F.; Guidoni, L.; Sinicropi, A. *J. Chem. Theory Comput.* **2012**, *8*, 112–124. (d) Jonasson, G.; Teuler, J.-M.; Vallverdu, G.; Mérola, F.; Ridard, J.; Lévy, B.; Demachy, I. *J. Chem. Theory Comput.* **2011**, *7*, 1990–1997. (e) Schaefer, L. V.; Groenhof, G.; Boggio-Pasqua, M.; Robb, M.A.; Grubmueller, H. *PLoS Comput. Biol.* **2008**, *4*, e1000034. (f) Virshup, A. M.; Punwong, C.; Pogorelov, T. V.; Lindquist, B. A.; Ko, C.; Martínez, T. J. *J. Phys. Chem. B* **2009**, *113*, 3280–3291. (g) Vendrell, O.; Gelabert, R.; Moreno, M.; Lluch, J. M. *J. Am. Chem Soc*. **2006**, *128*, 3564–3574. (h) Olsen, S.; Lamothe, K.; Martinez, T. J. *J. Am. Chem. Soc*. **2010**, *132*, 1192-1193.

(15) (a) Li, X.; Chung, L. W.; Mizuno, H.; Miyawaki, A.; Morokuma, K. *J. Phys. Chem. B* **2010**, *114*, 1114-1126. (b) Li, X.; Chung, L. W.; Mizuno, H.; Miyawaki, A.; Morokuma, K. *J. Phys. Chem. B* **2010**, *114*, 16666-16675. (c) Li, X.; Chung, L. W.; Mizuno, H.; Miyawaki, A.; Morokuma, K. *J. Phys. Chem. Lett.* **2010** *1*, 3328-3333. (d) Li, X.; Chung, L. W. Morokuma, K. in "*Computational Methods for Large Systems: Electronic Structure Approaches for Biotechnology and Nanotechnology*", ed. J. R. Reimers, John Wiley & Sons, Inc., Hoboken, NJ, 2011, pp. 397-433 and ref. 16j.

(16) (a) Maseras, F.; Morokuma, K. *J. Comput. Chem.* **1995**, *16*, 1170-1179. (b) Humbel, S.; Sieber, S.; Morokuma, K. *J. Chem. Phys.* **1996**, *105*, 1959-1967. (c) Matsubara, T.; Sieber, S.; Morokuma, K. *Int. J. Quantum. Chem*. **1996**, *60*, 1101-1109. (d) Svensson, M.; Humbel, S.; Froese, R. D. J.; Matsubara, T.;

Sieber, S.; Morokuma, K. *J. Phys. Chem.* **1996**, *100*, 19357-19363. (e) Svensson, M.; Humbel, S.; Morokuma, K. *J. Chem. Phys.* **1996**, *105*, 3654-3661. (f) Dapprich, S.; Komaromi, I.; Byun, K. S.; Morokuma, K.; Frisch, M. J. *J. Mol. Struc- THEOCHEM.* **1999**, *461-462*, 1-21. (g) Vreven, T.; Morokuma, K. *J. Comput. Chem.* **2000**, *21*, 1419-1432. (h) Vreven, T.; Byun, K. S.; Komaromi, I.; Dapprich, S.; Montgomery, J. A.; Morokuma, K.; Frisch, M. J. *J. Chem. Theory. Comput.* **2006**, *2*, 815-826. (i) Tao, P.; Fisher, J. F.; Shi, Q.; Vreven, T.; Mobashery, S.; Schlegel, H. B. *Biochemistry* **2009**, *48*, 9839-9847. (j) Chung, L. W.; Hirao, H.; Li, X.; Morokuma, K. *WIREs Comput. Mol. Sci.* **2012**, *2*, 327–350.

(17) Dolinsky, T. J.; Nielsen, J. E.; McCammon, J. A.; Baker, N. A. *Nucleic Acids Res.* **2004**, *32*, 665-667.

(18) Li, H.; Robertson, A. D.; Jensen, J. H. *Proteins.* **2005**, *61*, 704-721.

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

(20) (a) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926-935. (b) 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*, 5179-5197.

(21) Frisch, M. J. et.al, Gaussian 09 Revision **A.2.;** Gaussian, Inc.: Wallingford CT, **2009**.

(22) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. *J. Mol. Graph. Model.* **2006**, *25*, 247-260.

(23) Humphrey, W.; Dalke, A. and Schulten, K., "VMD - Visual Molecular Dynamics", J. Molec. Graphics, **1996**, *14*, 33-38.

(24) http://www.chemcraftprog.com.

(25) Andersson, K. et.al. Molcas Version 7.4, University of Lund: Sweden, **2004**.

(26) Ding, L. N.; Shen, L.; Chen, X. B.; Fang, W. H. *J. Org. Chem.* **2009**, *74*, 8956-8962.

(27) Yanai, T.; Tew, D. P.; Handy, N. C. *Chem. Phys. Lett.* **2004**, *393*, 51-57.

(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*, 259-267. (c) The choice of about six states in SA-CASSCF calculation: the IntraCT state with the largest *f* value was found to be of S_4 at this CASSCF(10e/7o) level. Therefore, at least five states should be included in the SA-CASSCF calculations. However, at several points, other state came into the S_4 or S_3 at the CASSCF level (please see **Table S7**). So, we have to consider at least 6 states in the states-average calculations to keep the initial five states consistently.

(29) Roos, B. O.; Andersson, K. *Chem. Phys. Lett.* **1995**, *245*, 215-223.

(30) Ghigo, G.; Roos, B. O.; Malmqvist, P.-A. *Chem. Phys. Lett.* **2004**, *396*, 142-149.

(31) Dapprich, S. K. r., I.; Byun, S.; Morokuma, K.; Frisch, M. *J. Mol. Struct.(Theochem)* **1999**, 462, 1-21.

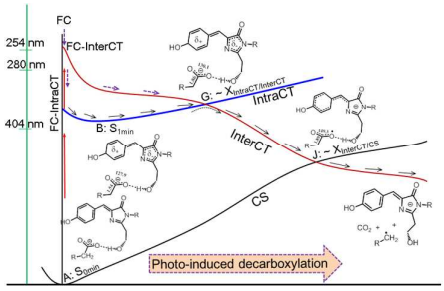
(32) (a) Field, M. J.; Bash, P. A.; Karplus, M. *J. Comput. Chem.* **1990**, *11*, 700-733. (b) Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1986**, *7*, 718-730.

(33) Altun, A.; Shaik, S.; Thiel, W. *J. Comput. Chem.* **2006**, *27*, 1324-1337.

(34) It is generally believed that excitation to higher excited states followed by ultra-fast internal conversion can give vibrationally "hot" S_1 due to transformation of some potential energy into thermal energy. However, quantum yield of the decarboxylation is low, it is hard to experimentally override this possibility at this moment and such possible mechanism has not been suggested.

(35) Roy et al. observed photo-induced decarboxylation in IrisFP. The mechanism was suggested to involve a biradical charge-transfer state (Kolbe-like mechanism) and relate to the formation of the distorted state (ref. 13a.). The difference between PA-GFP and IrisFP is possibly due to different protein environments and chromophore.

TOC drawing


FC
254 nm
280 nm
404 nm
FC-InterCT
FC-IntraCT
B: S1min
G: ~ XIntraCT/InterCT
IntraCT
InterCT
J: ~ XInterCT/ICS
ICS
A: S0min
Photo-induced decarboxylation