JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Article
pubs.acs.org/JACS

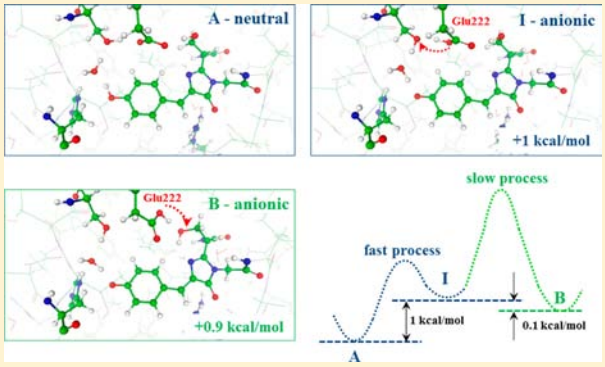
First-Principles Characterization of the Energy Landscape and Optical Spectra of Green Fluorescent Protein along the A→I→B Proton Transfer Route

Bella L. Grigorenko,[†,‡] Alexander V. Nemukhin,*[,†,‡] Igor V. Polyakov,[†] Dmitry I. Morozov,[†] and Anna I. Krylov[§]

[†]Chemistry Department, M.V. Lomonosov Moscow State University, Leninskie Gory 1/3, Moscow, 119991, Russian Federation
[‡]N.M. Emanuel Institute of Biochemical Physics, Russian Academy of Sciences, Kosygina 4, Moscow, 119334, Russian Federation
[§]Department of Chemistry, University of Southern California, Los Angeles, California 90089-0482, United States

Supporting Information

ABSTRACT: Structures and optical spectra of the green fluorescent protein (GFP) forms along the proton transfer route A→I→B are characterized by first-principles calculations. We show that in the ground electronic state the structure representing the wild-type (wt) GFP with the neutral chromophore (A-form) is lowest in energy, whereas the systems with the anionic chromophore (B- and I-forms) are about 1 kcal/mol higher. In the S65T mutant, the structures with the anionic chromophore are significantly lower in energy than the systems with the neutral chromophore. The role of the nearby amino acid residues in the chromophore-containing pocket is re-examined. Calculations reveal that the structural differences between the I- and B-forms (the former has a slightly red-shifted absorption relative to the latter) are based not on the Thr203 orientation, but on the Glu222 position. In the case of wt-GFP, the hydrogen bond between the chromophore and the His148 residue stabilizes the structures with the deprotonated phenolic ring in the I- and B-forms. In the S65T mutant, concerted contributions from the His148 and Thr203 residues are responsible for a considerable energy gap between the lowest energy structure of the B type with the anionic chromophore from other structures.



INTRODUCTION

The discovery and wide applications of the fluoroproteins of the green fluorescent protein (GFP) family inspired numerous investigations of the structure, dynamics, and spectra of these fascinating macromolecules. The results of these studies are summarized in several comprehensive reviews[1−8] as well as original reports. Despite significant coverage in the literature, many details regarding properties and functions of fluorescent proteins remain hidden, even those of the parent member of the family, the wild-type (wt) GFP itself. A complete atomic-level understanding of the structure and function of GFP is required to form a solid basis for rational design of novel efficient biomarkers for *in vivo* applications.

The current views on the GFP photocycle based on the canonical three-state model initially formulated in refs 9 and 10 are summarized in Scheme 1.

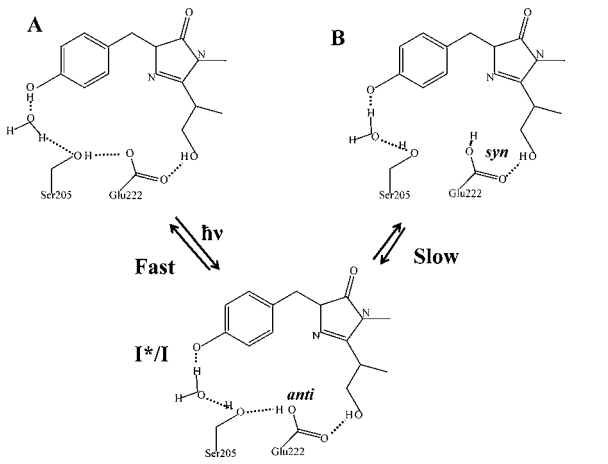
One of the conformational forms of wt-GFP (termed A) is characterized by the 395 nm (3.14 eV) absorption band maximum; in this form the chromophore is neutral (that is, protonated at the phenolic oxygen) and the nearby Glu222 residue is deprotonated. Another conformational form, termed B, absorbs at 475 nm (2.61 eV); it features the anionic (phenolate) form of the chromophore and protonated Glu222. Under physiological conditions, the ground-state population of form B is approximately one-sixth that of form A, suggesting that in the ground state A is lower in energy.[9] Fluorescence at 508 nm (2.44 eV) is assigned to the forms with the anionic chromophore. Transformations between A and B occur through an intermediate termed I. Upon excitation of the neutral chromophore, a series of coupled proton-transfer steps occur through a hydrogen-bonding network that includes the phenolic ring of the chromophore, the nearby water molecule, and the Ser205 and Glu222 side chains, giving rise to the deprotonated (anionic) chromophore and protonated Glu222. The change from A to I is solely a protonation change, whereas the change from I to B is a slow conformational change.

The crystal structure of wt-GFP as a dimer was solved at 1.9 Å resolution and deposited to the Protein Data Bank (PDB) with the code 1GFL.[11] wt-GFP shows dual-wavelength (395 and 475 nm) absorption, revealing the coexistence of A and B conformational forms. Mutation S65T suppresses the 395 nm absorption band while amplifying and slightly shifting (from 475 to 489 nm)[10] the longer-wavelength absorption peak. The

Received: March 9, 2013
Published: July 9, 2013

dx.doi.org/10.1021/ja402472y | J. Am. Chem. Soc. 2013, 135, 11541−11549

Scheme 1. Main Features of the Three-State Model[9,10] of Conformational Changes in the GFP Photocycle



emission maximum is almost the same, i.e., 511 nm in S65T vs 508 nm in wt-GFP. Consequently, the structure of the S65T GFP mutant determined at 1.9 Å resolution (PDB code 1EMA[12]) is often considered as a good representation of form B. The monomeric wt-GFP was also crystallized, and the structure was determined (PDB code 1EMB,[10] resolution 2.1 Å) by molecular replacement on the basis of the S65T model. Other crystal structures of GFP mutants with the anionic and neutral chromophores have also been reported, e.g., refs 13−15, including that (PDB code 2WUR[15]) of the highest resolution of 0.9 Å.

Importantly, no crystal structure is available for the I-form denoting the species absorbing at the most red-shifted wavelengths (∼495 nm). However, spectroscopic investigations[16−23] of the GFP photocycle have been interpreted in terms of proton transfer routes that may involve several intermediates with the anionic chromophore, presumably, of the I-type, in the electronic ground state.

From the theoretical side, Lill and Helms applied classical molecular dynamics (MD) augmented by instantaneous proton hopping to simulate the ground-state structures along the proton shuttle route.[24] On the basis of the MD simulations, the authors suggested the existence of two I-type conformations, a "nonrelaxed" form with the side chain of Glu222 in an *anti* orientation and a "relaxed" form with the *syn* orientation (see Scheme 1). Concerning the conformational changes along the I→B transformation, some evidence favoring internal rotation of the Thr203 side chain was provided by molecular mechanical[25] and classical MD[26] simulations. It should be noted, however, that the hypothesis about a stabilizing effect of the Thr203 side-chain rotation on the emerging negative charge of the phenolic oxygen following the proton transfer during the A→B transformation was formulated earlier,[10] when analyzing the crystal structures of GFP.

Surprisingly, no reliable quantum mechanical calculations of the energy landscape along the A→I→B route have been reported so far. There are a few quantum chemical studies of considerably simplified model systems. The model reported by Zhang et al.[27] consisted of molecular groups mimicking the GFP chromophore, the side chains of Ser205, Glu222, and His148, and the water molecule in the gas phase and in the dielectric continuum. Using density functional theory (DFT) with the B3LYP functional, they reported that their representation of the I-form is 10−30 kcal/mol lower than their A-form, which is far too large a gap.[27] Vendrell et al.[28,29] employed complete active space self-consistent field (CASSCF) with the perturbation corrections (CASPT2) for an almost identical model system (without His148 side) in the gas phase. They arrived at an opposite conclusion, namely, that the I-form is 18 kcal/mol higher in energy than the A-form.[28] Owing to an oversimplified model system used in both studies,[27,28] the computed energy differences (±20 kcal/mol) between the structures with the neutral and anionic chromophores can hardly be related to the actual protein energetics.

Calculations of the electronic spectra of the bare (see refs 2, 5, 7, 30, 31 and references therein) or protein-bound GFP chromophore are more frequent in the literature.[28,29,32−39] Here we focus only on the quantum mechanical/molecular mechanical (QM/MM)[40] studies that include the effect of the protein. Olivucci and co-workers[33] used the CASSCF-based methods (QM) and the CHARMM force field (MM) to estimate the positions of the $S_0 \rightarrow S_1$ and $S_1 \rightarrow S_0$ transitions in GFP with the anionic chromophore. The computed values agreed well (within 20−30 nm) with the experimental band maxima. By considering a series of models with gradually increasing quantum subsystems, the authors investigated, in particular, the effect of the charged amino acid residue Arg96 located near the chromophore on the computed electronic spectrum. Bravaya et al.[36] characterized excited and ionized states of enhanced GFP using advanced electronic structure methods within the QM/MM scheme; they reported good agreement with experiment and found that the overall effect of the protein environment on excitation energies is small. Optical spectra of GFP and several mutated variants with different protonation states of the chromophore were calculated by Hasegawa et al.[34] using a configuration-interaction-based method for electronic energy differences between the ground and excited states and accounted for the effect of the protein matrix by QM/MM. A reasonable agreement between the calculated and experimental excitation/emission energies was reported. Filippi et al.[37] computed vertical excitation energies of the neutral A- and anionic B-forms of wt-GFP using a variety of electronic structure methods applied to the protein model constructed using QM/MM MD simulations. Electronically excited states of the model system representing the neutral GFP chromophore in the binding pocket of the protein were characterized by Grigorenko et al.[38] by using advanced quantum chemistry methods. The calculations allowed the authors to identify the states that have Glu222→Chro charge-transfer character and to determine whether such states (proposed as gateway states leading to decarboxilation of nearby Glu) have sufficient oscillator strengths and excitation energies to be accessible using UV or visible light, either directly or in a two-step process. A subsequent paper by Morokuma et al.[39] also investigated higher excited states in GFP in the context of photoinduced decarboxylation using the ONIOM version of QM/MM.

The present study characterizes the energy landscape along the A→I→B route using reliable quantum chemical methods combined with the QM/MM representation of the protein, which is the first attempt of this kind. We construct a series of molecular models describing the photoinduced intermediates for wt-GFP and the S65T mutant consistent with the available structural and spectroscopic experimental data. We combined an intensive search of the minimum energy configurations in the ground S_0 and excited S_1 electronic states with high-level electronic structure calculations within the QM/MM scheme.

dx.doi.org/10.1021/ja402472y | J. Am. Chem. Soc. 2013, 135, 11541−11549

These calculations allowed us to pinpoint relative energetics of different structural forms of GFP. For the first time, the three-state model conjecture based on experimental evidence has received concrete theoretical support. The calculations also allowed us to re-examine the role of the nearby amino acid residues in the chromophore-containing pocket.

MODELS AND METHODS

The initial set of model systems was based on the coordinates of heavy atoms from the PDB entry 1EMA,[12] i.e., the S65T mutant of GFP. We restored the side chain of Ser at position 65 and added hydrogen atoms by assigning customary protonation states of polar residues (Lys, Arg, Glu, Asp). Water molecules resolved in the crystal structure were also included in the model systems. To obtain minimum energy configurations, we applied the flexible effective fragment version[41,42] of QM/MM based on the effective fragment potential approach.[43,44] In this scheme, groups assigned to the MM part are represented by effective fragments contributing their one-electron potentials to the quantum Hamiltonian;[43] the peptide chains of the protein are described as flexible chains of small effective fragments,[41,42] and fragment−fragment interactions are computed with conventional force fields. A modified code of the GAMESS(US)[45,46] program and TINKER[47] were used in the QM/MM calculations.

The QM part included the chromophore (Chr), the side chains of Arg96, Glu222, Ser205, His148, and Thr203, and two water molecules. Ground-state (S_0) structures were optimized by using DFT with the PBE0 functional,[48] which was shown to perform well in geometry optimizations of organic molecules,[49] and the 6-31G* basis set. Energies and forces in the MM subsystem were computed using the AMBER force field parameters. Analysis of the hydrogen-bond network around the chromophore allowed us to suggest several arrangements of the participating molecular groups. In particular, we noticed different possible orientations of the side chains of Thr203, His148, and Glu222. In sum, we located several minimum energy structures within 10 kcal/mol range for the system representing wt-GFP with the anionic chromophore. Several lowest energy structures will be discussed in the forthcoming section. For S65T, we restricted our analysis by three minimum energy conformations with the anionic chromophore.

For both systems (wt-GFP and the S65T mutant) we prepared the structures with the neutral chromophore by transferring manually the protons between the chromophore and Glu222, followed by complete QM/MM optimizations. Importantly, this computational protocol enables us to compare directly the energies of the structures representing conformations of GFP with the neutral and anionic chromophores. In particular, for the system representing wt-GFP, the structure with the neutral chromophore (presumably, A-form) was found to be 0.9 kcal/mol lower in energy than the lowest energy structure with the anionic chromophore (presumably, B-form). According to the Boltzmann formula, this energy gap leads to 5:1 ratio between the two protonated states, which is in a remarkable agreement with an experimentally derived 6:1 population ratio.[9] In contrast to wt-GFP, in the S65T mutant our calculations show that the system with the anionic chromophore lies much lower in energy that the system with the neutral chromophore. This result is also in line with the experimental data.

Vertical excitation energies at all located stationary points on the ground-state potential energy surface (PES) were computed using eXtended Multi-Configurational Quasi-Degenerate Perturbation Theory (XMCQDPT2),[50] an advanced quantum chemical method for excited states of organic fluorophores (e.g., see ref 51). The XMCQDPT2 calculations were carried out using the CASSCF(12/11)/cc-pVDZ wave functions for the clusters composed of Chr, Arg, Glu, Ser, and water embedded inside the set of effective fragments representing the rest of the model system. The Firefly program[52] was used in these calculations.

To compute the emission spectra, we located minimum energy points on the excited-state (S_1) PES using the configuration interaction singles (CIS) method. Vertical $S_1 \rightarrow S_0$ transition energies were calculated using the same approach as for the excitation energies. We expect that the accuracy of the estimated emission energies is somewhat lower than that of the excitation energies.

To further validate the computed energy differences, we performed additional calculations using another *ab initio* approach based on the scaled-opposite-spin second-order corrected configuration interaction singles, or SOS-CIS(D).[53,54] The SOS-CIS(D)/cc-pVDZ calculations were performed with the Q-Chem program[55] for the QM parts from the corresponding QM/MM model systems. As reported below, both computational approaches resulted in quantitatively consistent results.

RESULTS AND DISCUSSION

Structure of wt-GFP with the Neutral Chromophore, Neutral-wt-A1. We begin by considering the lowest-energy structure of wt-GFP containing the neutral chromophore found in our calculations. As explained above, the model system with the neutral chromophore was prepared manually by assigning protons to molecular groups of Chr, Ser205, and Glu222 starting from the QM/MM optimized system with the anionic chromophore (namely, from the structure **anion-wt-I1** described below) and then completely reoptimized using QM/MM.

Figure 1 shows the equilibrium structure of the fragment of the chromophore-containing pocket that includes side chains

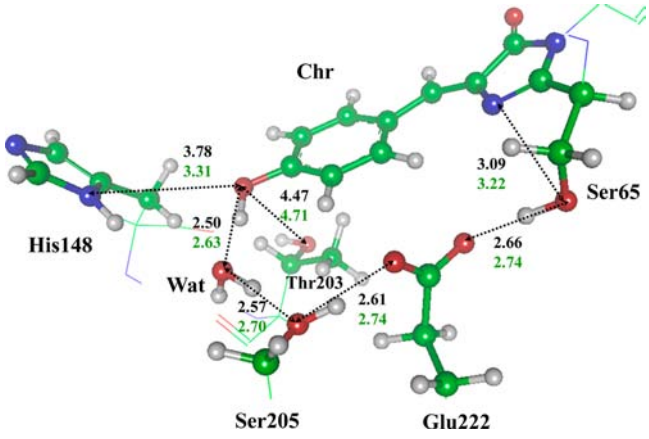


Figure 1. Fragment of the model system representing wt-GFP with the neutral chromophore, **neutral-wt-A1**. Here and in all figures, carbon atoms are shown in green, oxygen in red, and nitrogen in blue. Distances (in Å) between heavy atoms for the QM/MM optimized structures (top) and for the crystal structure 1GFL[11] (bottom) are shown.

His148, Thr203, Ser205, Glu222, and water. The side chain of Arg92, which is also included to the QM subsystem, is omitted for clarity.

Figure 1 compares several computed structural parameters, namely, the distances between heavy atoms, illustrating hydrogen bond patterns, to those from crystal structure 1GFL[11] that can be considered as a good representation of the A-form with the neutral chromophore. Our computed geometric parameters (Figure 1) agree well with those from the crystal structure, within usual discrepancies between the QM/MM optimized structures and the values deduced directly from X-ray data.[56] Note that our starting coordinates were taken from a different crystal structure, 1EMA,[12] corresponding to the S65T mutant of GFP with the anionic chromophore; the coordinates were then reoptimized after a proton was added to the phenolic oxygen. Thus, comparison of the computed

dx.doi.org/10.1021/ja402472y | J. Am. Chem. Soc. 2013, 135, 11541−11549

structure with 1GFL presents a stringent test for our computational protocol.

The computed vertical S_0-S_1 excitation energy is 3.19 eV (388 nm) and 3.18 eV (390 nm) when using the XMCQDPT2 and SOS-CIS(D) methods, respectively. These values are in excellent agreement with the experimental band maximum at 395 nm for the A-form of wt-GFP.[9,10]

Starting from the ground-state coordinates, we optimized the structure of this model system in the first excited state (S_1) by using CIS. The shape of the S_1 PES features a very shallow minimum in the Franck–Condon region, and the barriers separating it from the downward-sloping valleys are small. The computed S_1-S_0 energy gap is 2.68 eV (461 nm) at the XMCQDPT2 level and 2.73 eV (455 nm) using SOS-CIS(D). Experiments[16,22] show weak emission bands in the 440–460 nm region that most likely belong to the A-form.

On the basis of the agreement between the computed and experimental spectroscopic and structural data, we believe that this structure, termed **neutral-wt-A1** (Figure 1), represents accurately the A-form of wt-GFP. Importantly, it has the lowest energy among all model structures for wt-GFP considered in this work. Below, the energies of structures of wt-GFP with anionic chormophores in the ground state will be given relative to the energy of this **neutral-wt-A1** structure.

Structure of wt-GFP with the Anionic Chromophore, Anion-wt-I1. Now we return to the chronologically first model system developed in this work, namely, the system with the anionic chromophore designed following the crystal structure 1EMA[12] in which the side chain of Ser at position 65 was restored. The chromophore is in the phenolate form (i.e., deprotonated at the phenolic oxygen), but the Glu222 side chain is protonated and oriented in the *anti* position. A fragment of the QM/MM optimized structure in the chromophore-containing pocket is shown in Figure 2 from the same angle as the **neutral-wt-A1** structure. In this figure we compare critical distances between heavy atoms, illustrating hydrogen-bond patterns in the three structures, i.e., (i) one obtained in the QM/MM optimization (top), (ii) the 1EMB[10] crystal structure presumably representing wt-GFP with the anionic chromophore (middle), and (iii) the structure obtained in the QM/MM optimization of the system with the neutral chromophore, **neutral-wt-A1** (bottom). The energy of this system is 1.0 kcal/mol above that of **neutral-wt-A1**.

Comparison of QM/MM equilibrium structure of this system with the anionic chromophore (Figure 2) with **neutral-wt-A1** (Figure 1) reveals that no dramatic changes in the hydrogen bond network occur due to the concerted proton transfer from one moiety to another.

This model system is a good candidate for the intermediate (I) form of wt-GFP; we will call it **anion-wt-I1**. The computed positions of the absorption and emission bands strongly support this assignment.

The computed vertical S_0-S_1 excitation energy is 2.52 eV (492 nm) at both XMCQDPT2 and SOS-CIS(D) levels of theory. This value almost coincides with the experimental absorption band at 495 nm determined in hole burning experiments.[16] After locating the energy minimum on the excited-state PES, we computed the S_1-S_0 transition energy to be 2.37 eV (523 nm) at the XMCQDPT2 level (the respective SOS-CIS(D) value is 533 nm). The experimental emission band maximum in wt-GFP assigned to the I-form is 508 nm (2.44 eV);[9,10] the discrepancies of ~0.1 eV in emission are within the anticipated error bars of our computational protocol.

Lowest Energy Structure of wt-GFP with the Anionic Chromophore, Anion-wt-B1. An extensive search of minimum energy structures with the anionic chromophore on the ground electronic state PES allowed us to locate a structure with energy slightly lower (by 0.1 kcal/mol) than that of the previously considered **anion-wt-I1**. This structure, here termed **anion-wt-B1**, with energy 0.9 kcal/mol above that of **neutral-wt-A1**, is shown in Figure 3. We note a different orientation of the Glu222 side chain, which assumes the *syn* position in this structure.

When comparing selected geometric parameters computed for **anion-wt-B1** to those resolved in the crystal structure 1EMB,[10] we note a good correlation in the Chr-Thr203-His148-Wat region and some discrepancies in the region Wat-Ser205-Glu222-Ser65. Figure 3 also compares geometric parameters of the two computationally designed structures,

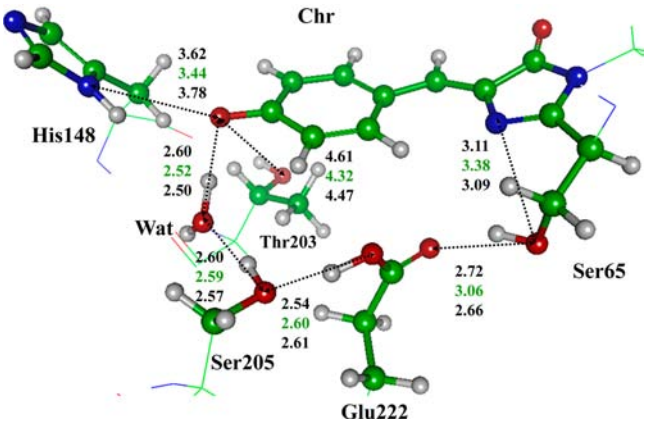


Figure 2. Fragment of the QM/MM optimized structure with the anionic chromophore, **anion-wt-I1**, located 1.0 kcal/mol above the **neutral-wt-A1** structure. Distances (in Å) between heavy atoms are arranged as follows: the results of our QM/MM optimization (top), the values from the crystal structure 1EMB[10] (middle), and the values from the QM/MM optimized structure with the neutral chromophore from Figure 1 (bottom).

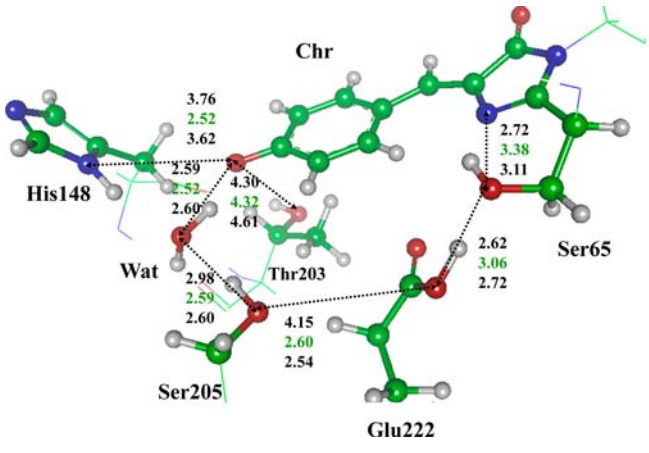


Figure 3. Fragment of the QM/MM optimized structure with the anionic chromophore, **anion-wt-B1**, with energy 0.9 kcal/mol above that of **neutral-wt-A1**. Distances (in Å) between heavy atoms are arranged as follows: the results of the QM/MM optimization (top), the values from the crystal structure 1EMB[10] (middle), and the values from the QM/MM optimized structure, **anion-wt-I1**, from Figure 2 (bottom).

dx.doi.org/10.1021/ja402472y | J. Am. Chem. Soc. 2013, 135, 11541–11549

anion-wt-B1 and **anion-wt-I1**. Apparently, the hydrogen-bond pathway responsible for the proton shuttle Chr-Wat-Ser205-Glu222-Ser65 is disrupted in **anion-wt-B1**. Therefore, in contrast to **anion-wt-I1** (Figure 2), wihch may be efficiently involved in the forward and backward proton shuttle between the states with the neutral and anionic chromophores, in the case of **anion-wt-B1**, the system once falling in this region will be trapped there, and the backward proton transfer would require considerable energy to properly rearrange the side chain of Glu222 from *syn* to *anti* position. We estimated the potential barrier that separates the **anion-wt-B1** and **anion-wt-I1** structures by computing a series of minimum energy points along the rotational coordinate. A point where the energy gradient with respect to this coordinate changed its sign was taken as the saddle point, with energy of 12 kcal/mol. This relatively high value is consistent with the experimental evidence on the "slow or infrequent" transition from I to B.[10]

The computed S_0-S_1 vertical excitation energy is 2.58 eV (484 nm) using XMCQDPT2; the SOS-CIS(D) value is 2.54 eV (489 nm). In accord with the experimental data, these values are slightly blue-shifted with respect to those of **anion-wt-I1**. The computed emission band maxima, 519 and 536 nm using XMCQDPT2 and SOS-CIS(D), respectively, are within our error bars (~0.1 eV) from those of **anion-wt-I1** (523 and 533 nm).

Thus, this model system, **anion-wt-B1**, is a good representation of the canonical B-form of wt-GFP.

Higher Energy Conformations of wt-GFP with the Anionic Chromophore. In addition to model structures **anion-wt-B1** and **anion-wt-I1**, we were able to locate several more local minima on the ground-state PES of the model system with the anionic chromophore. We analyzed a hypothesis traced to the pioneer studies of GFP[9,10] that, upon the I→B transformation, the side chain of Thr203 should rotate to stabilize the emerging negative charge at the phenolic oxygen. Therefore, we prepared manually a new model system by rotating the side chain of Thr203 starting from the **anion-wt-B1** structure and reoptimized its coordinates in the QM/MM calculation.

Figure 4 shows the **anion-wt-B1** structure (from Figure 3) in the bottom panel and the newly optimized structure, **anion-wt-B2**, in the top panel. Both of them contain the side chain of Glu222 in the *syn* orientation with the disrupted proton transfer wire. The system shown in the top panel demonstrates that the side chain of Thr203 is turned toward the chromophore's phenolic oxygen; however, this structure (called **anion-wt-B2**) is 3.1 kcal/mol higher in energy than **anion-wt-B1**, in which Thr203 is located farther away from the chromophore. We also note that the distance between the phenolic oxygen and Nδ of His148 is considerably shorter in **anion-wt-B1** (3.76 Å) than in **anion-wt-B2** (4.12 Å). It appears that Chr-His148 interaction is more important for stabilization of the B-form of wt-GFP than Chr-Thr203.

To locate another minimum energy configuration, we manually rotated the side chain of Thr203 from the position in **anion-wt-I1** to approach the chromophore closer and reoptimized the structure by QM/MM. The resulting model structure (which we call **anion-wt-I2**) is shown in the middle panel in Figure 5. It is similar to Figure 4, with the major difference being the orientation of the Glu222 side chain: in the **anion-wt-I2** structure (as well as in **anion-wt-I1**) the hydrogen-bond pattern Chr-Wat-Ser205-Glu222-Ser65 permits proton transfer, but in the **anion-wt-B1** and **anion-wt-B2** structures it does not. The energy of **anion-wt-I2** is 3.0 kcal/mol, that is, 2.0 kcal/mol higher than that of **anion-wt-I1**. Therefore, comparison of the energies in the pairs of structures, **anion-wt-B1** (Thr203-out) vs **anion-wt-B2** (Thr203-in), and **anion-wt-I1** (Thr203-out) vs **anion-wt-I2** (Thr203-in), shows that turning the side chain of Thr203 to the position closer to the chromophore does not lead to energy lowering; the structures with Thr203 in the outward position are, in fact, more stable by 2−3 kcal/mol.

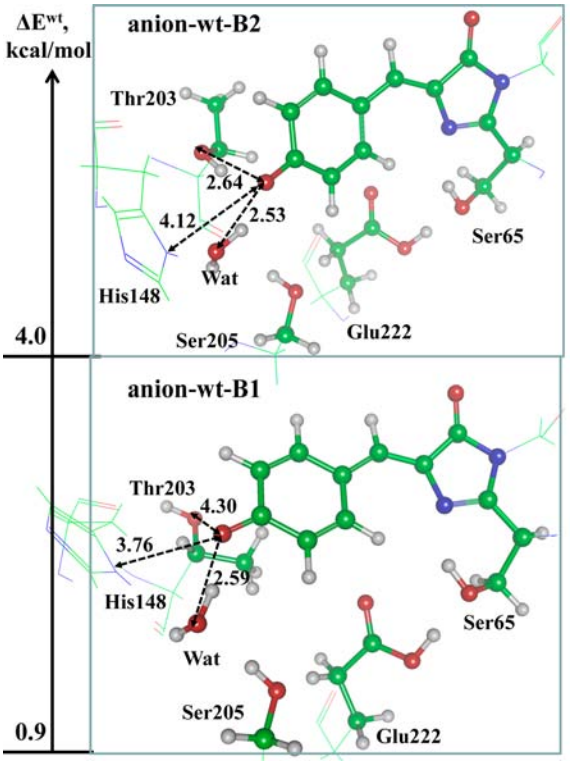


Figure 4. Fragments of the QM/MM optimized structures with the anionic chromophore, **anion-wt-B1** (below) and **anion-wt-B2** (above). Distances between heavy atoms are given in Å.

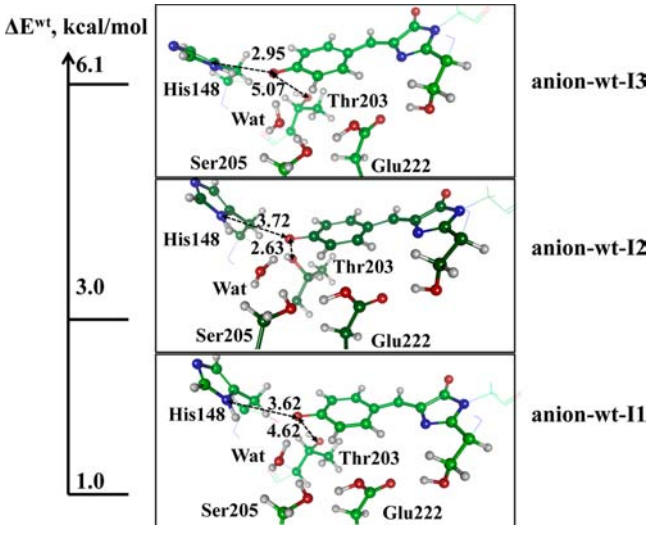


Figure 5. Structures and energies of the three wt-GFP model systems with the anionic chromophore of the I type. Distances (in Å) from oxygen of the chromophore to Nδ of His148 and Oγ of Thr203 are given.

dx.doi.org/10.1021/ja402472y | J. Am. Chem. Soc. 2013, 135, 11541−11549

It seems that the interaction of the deprotonated chromophore with the His148 side chain provides sufficient stabilization of the anionic chromophore even without the involvement of Thr203. Importantly, the position of the His148 side chain is fairly flexible. We located a conformation called **anion-wt-I3** with the energy of 6.1 kcal/mol. This structure shown in the top panel in Figure 5 almost coincides with the structure **anion-wt-I1**; the major difference is in the position and orientation of the side chain of His148 relative to the chromophore. We mention this conformation first of all to demonstrate that, using computational chemistry techniques, one can indeed find different stable and metastable conformers beyond the canonical forms A, I, and B, as suggested in several experimental studies,[17−23] and, second, to demonstrate the sensitivity of the protein properties to fairly small conformational changes. Here, a slight shift of one residue, His148, rearranging its hydrogen bond patterns results in a noticeable change in energy by about 5 kcal/mol.

With respect to the spectral bands, the considered variations in both types of structures (I-like and B-like) result in almost negligible changes: the absorption bands in **anion-wt-B1** and **anion-wt-B2** are within 2 nm from 484 nm, whereas the absorption bands in **anion-wt-I1**, **anion-wt-I2**, and **anion-wt-I3** are within 2 nm from 492 nm.

Structures of the S65T Mutant with the Anionic Chromophore. Figure 6 shows the QM/MM optimized structure of the lowest energy model system with the anionic chromophore containing Thr at position 65. It is important to note that the energy of this structure is considerably lower than the energy of the S65T model with the neutral chromophore, as discussed below. This result is in agreement with the experimental observation that the spectra of the S65T mutant are similar to the B-form of GFP and have no bands corresponding to the A-form.[10] Below, we will report the energies (ΔE^{S65T}) of other structures representing the S65T mutant with respect to this lowest energy structure.

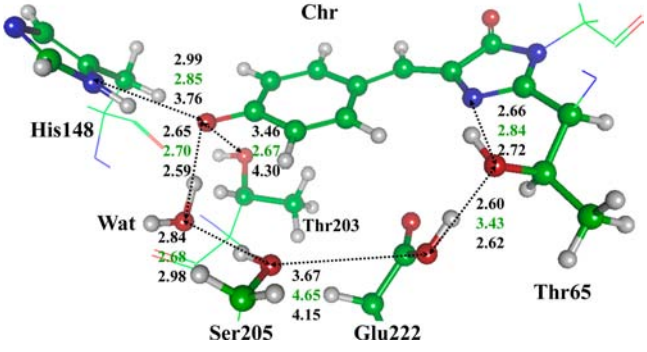


Figure 6. Fragment of the QM/MM optimized structure of the S65T mutant with the anionic chromophore, **anion-S65T-B1**. Distances (in Å) between heavy atoms are arranged as follows: the results of this QM/MM optimization (top), the values from the crystal structure, 1EMA,[12] (middle), and the values from QM/MM optimized structure, **anion-wt-B1**, from Figure 3 (bottom).

When discussing wt-GFP with the anionic chromophore, we could distinguish the I-series, **anion-wt-I1** (Figure 2, bottom panel in Figure 5), **anion-wt-I2** (middle panel in Figure 5), and **anion-wt-I3** (top panel in Figure 5), in which the proton wire expands over the Chro-Wat-Ser205-Glu222 chain, and the B-series, **anion-wt-B1** (Figure 3) and **anion-wt-B2** (Figure 4), in which the proton wire is disrupted between Ser205 and Glu222. Clearly, the structure shown in Figure 6 belongs to the B-series; therefore, we call it **anion-S65T-B1**.

Figure 6 compares the critical distances between heavy atoms obtained in the QM/MM optimization (top), in the 1EMA[12] crystal structure (middle), and in **anion-wt-B1** mimicking the B-form of wt-GFP. Interestingly, both His148 and Thr203 side chains are noticeably closer to the phenolic oxygen of the chromophore in S65T than in wt-GFP (cf. Figure 6 and Figure 3), thus providing better stabilization of the S65T structure with the anionic chromophore. This is consistent with a large energy gap (between 5.2 and 17.9 kcal/mol) separating the **anion-S65T-B1** structure from other structures for the S65T mutant discussed below. In the case of wt-GFP, the lowest-energy anionic structure corresponding to the B-form, **anion-wt-B1** (Figure 3), lies 0.9 kcal/mol above **neutral-wt-A1** (A-form) and within 0.1 kcal/mol from the **anion-wt-I1** (I-form) structures. Thus, the wt-GFP exhibits a variety of low-energy conformers, whereas the S65T mutant is best represented by the B-form.

The computed vertical S_0-S_1 excitation energy for **anion-S65T-B1** structure is 2.50 eV (495 nm, XMCQDPT2); the SOS-CIS(D) value is 503 nm. Therefore, the calculated shift in the B-type absorption band due to the S65T mutation is +11 nm (484→495 nm), to be compared with the experimental value of +14 nm (475→489 nm).[10]

After locating the energy minimum on the excited-state (S_1) surface, we computed the S_1-S_0 transition energy of **anion-S65T-B1** to be 2.38 eV (519 nm); the SOS-CIS(D) value is 518 nm. This is practically the same as for the **anion-wt-B1** model structure, again, in almost perfect agreement with the experimental values when comparing the emission bands of wt-GFP and S65T mutant (508 and 511 nm).[9,10]

We located two more stationary points on the ground-state PES of the S65T model system with the anionic chromophore (Figure 7). One of them, lying 5.2 kcal/mol above the lowest energy minimum, belongs to the I-series (**anion-S65T-I1**), while another (at 7.3 kcal/mol) belongs to the B-series (**anion-**

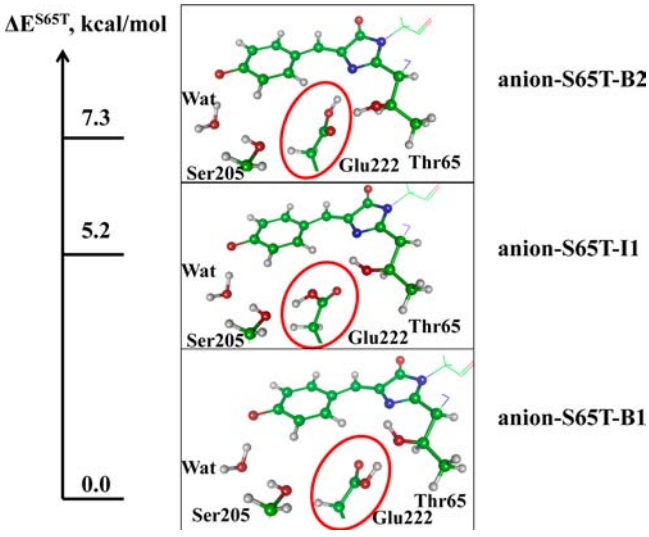


Figure 7. Structures and energies of the three S65T model systems with the anionic chromophore. Orientation of the Glu222 side chain is different in these structures.

dx.doi.org/10.1021/ja402472y | J. Am. Chem. Soc. 2013, 135, 11541−11549

S65T-B2). Their absorption bands are 479 and 483 nm, respectively, at the XMCQDPT2 level of theory.

Structure of the S65T Mutant with the Neutral Chromophore. The model system mimicking the S65T mutant with the neutral chromophore, **neutral-S65T-A1**, is shown in Figure 8. Most importantly, the energy of this structure is too high (17.9 kcal/mol above **anion-S65T-B1**), making it irrelevant for the properties of the protein under normal conditions.

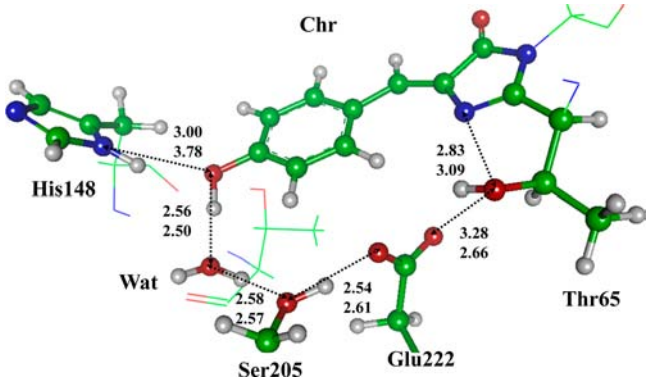


Figure 8. Fragment of the system representing the S65T mutant with the neutral chromophore. Distances (in Å) between heavy atoms are arranged as follows: the results of our QM/MM optimization for this structure (top) and the values from the QM/MM optimized structure representing wt-GFP from Figure 1 (bottom).

Excited-State Structures. The emphasis of our simulations was on the structures in the ground electronic state, S_0. The location of minimum energy configurations on the excited-state (S_1) PES performed using CIS allowed us to evaluate vertical transition energies roughly corresponding to the band positions in the emission spectra. Such calculations provide additional validation of the ground-state structures. To characterize more accurately structures in the excited state, higher level quantum chemistry approaches are required, e.g., see refs 57 and 58. Here we only briefly comment on issues relevant to our simulations.

It is well known[31] that electronic excitation leads to a moderate bond length increase (0.04−0.05 Å) in the two bridge CC bonds of the anionic chromophore. The structural changes in the neutral form of the chromophore are different: the formally single bond contracts by about 0.04 Å, whereas the formally double bond elongates by 0.08 Å. We observe these trends when comparing minimum-energy structures on the S_0 and S_1 PESs.

To illustrate changes in structures upon excitation, we show (Figure 9) several interatomic distances in the ground and excited states obtained in the QM/MM optimizations for wt-GFP with the neutral chromophore. Such changes are typical for all conformations considered in this work.

Similarly to the ground-state case, we located several minimum energy conformations on the excited-state surface. In particular, another structure was found for the excited-state system, **neutral-wt-A2***, which differs from **neutral-wt-A1*** shown in Figure 9 by a different position of the His148 side chain which moves closer to the chromophore. Energies of these two conformations on the S_1 PES were fairly similar. One may speculate that these two excited-state structures are responsible for features observed in spectroscopy experiments.

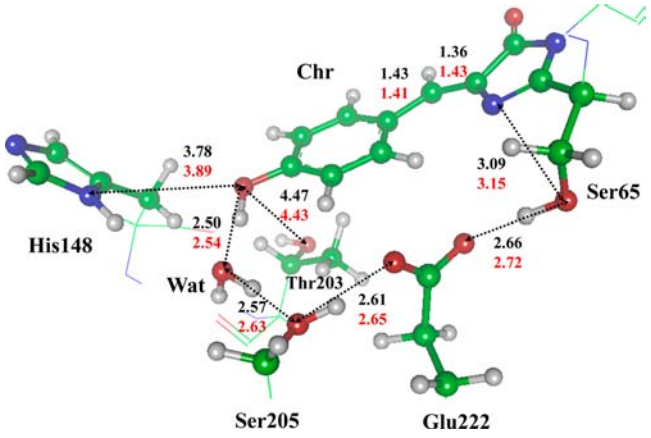


Figure 9. Comparison of selected geometry parameters of the system representing wt-GFP with the neutral chromophore in the ground state, **neutral-wt-A1** (top values in black), and in the excited state, **neutral-wt-A1*** (bottom values, red).

According to these studies,[20,22,23] other intermediates beyond the conventional I*-form are involved in the proton transfer process in GFP. The hypothesis formulated in ref 20 assumes an intermediate in which the proton wire "has partially adjusted to the excited state, but is not optimal for proton transfer yet". We attempted to locate the excited-state conformation similar to the chemical structures shown in Scheme 1 of ref 20; however, the calculations converged to the **neutral-wt-A2*** structure described above. Thus, although our results support the existence of multiple conformers, no partially proton-transferred structure has been found. As mentioned above, more accurate excited-state calculations are required to characterize dynamics of proton transfer.

Summary. To summarize the results obtained in this work, we collect in Table 1 the computed relative energies of major conformers referenced to the corresponding zero-energy structures: **neutral-wt-A1** for wt-GFP or **anion-S65T-B1** for the S65T mutant.

Table 1. Computed Relative Energies (kcal/mol) of the Ground-State Species for wt-GFP and S65T

wt-GFP		S65T	
neutral-wt-A1 (Figure 1)	0	anion-S65T-B1 (Figure 6)	0
anion-wt-B1 (Figure 3)	0.9	anion-S65T-I1 (Figure 7)	5.2
anion-wt-I1 (Figure 2)	1.0	anion-S65T-B2 (Figure 7)	7.3
anion-wt-I2 (Figure 5)	3.0	neutral-S65T-A1 (Figure 8)	17.9
anion-wt-B2 (Figure 4)	4.0		
anion-wt-I3 (Figure 5)	6.1		

Figure 10 schematically illustrates the computed energy landscape for wt-GFP. We estimated the energy barriers separating the major isomers on the ground-state potential surface in the QM/MM approach: the barrier between the A- and I-forms is very low, but the barrier between the I- and B-forms is about 12 kcal/mol.

Table 2 summarizes the spectral properties computed using XMCQDPT2 and SOS-CIS(D).

The anticipated accuracy in the excitation energy calculations in these photoactive proteins using protocols employed in our study is illustrated by the results of modeling several families of photoreceptor proteins, as summarized in Figure 1 of ref 59. The observed wavelengths can be reproduced within

dx.doi.org/10.1021/ja402472y | J. Am. Chem. Soc. 2013, 135, 11541−11549

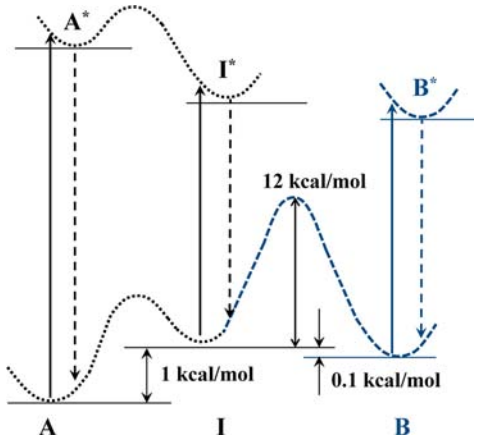


Figure 10. Computed energy diagram for the wt-GFP structural forms.

approximately 3 kcal/mol (~0.13 eV) error bar. Our results are consistent with these estimates.

CONCLUSIONS

This paper focuses on the QM/MM computed structures and spectra of the model systems mimicking the chromophore-containing domains of wt-GFP and the S65T mutant. The novel aspect of this study is that we are able to arrange different structures by energy. In particular, we show that the system with the neutral chromophore (form A) is the lowest in energy for wt-GFP, whereas the systems with the anionic chromophore that are assigned to the B- and I-forms are about 1 kcal/mol higher in energy. The respective population ratio between A and B (5:1) is in excellent agreement with the experimentally derived one (6:1).[9]

In the S65T mutant, the systems with the anionic chromophore are much lower in energy (about 18 kcal/mol) relative to the neutral chromophore structure.

Another important finding came from the re-examination of the role of the nearby amino acid residues in stabilizing the B-form with the anionic chromophore, which is produced from the A-form with the neutral chromophore upon excited-state proton transfer through the proton wire Wat-Ser205-Glu222. According to our calculations, the conformations with the rotated side chain of Thr203 (relative to its initial orientation in the A-form) do not lead to energy lowering. In the case of wt-GFP, the contributions from the His148 residue hydrogen-bonded to the chromophore are sufficient to stabilize the structures with the deprotonated phenolic chromophore's ring in both I- and B-forms. In the case of S65T, the concerted contributions from His148 and Thr203 residues lead to considerable energy separation between the lowest energy structure of the B type with the anionic chromophore from other structures. Correspondingly, we reassigned the I- and B-forms of GFP. In accord with the previous knowledge and our scheme, the I-forms refer to the species that absorb at wavelengths slightly red-shifted with respect to the B-form; however, the structural criteria that distinguish B and I species are not based on the Thr203 orientation, as was originally suggested.[10] Our results indicate that the orientation of the protonated Glu222 side chain is responsible for the difference between the I- and B-forms. In the I structures, the Glu222 side chain is in the *syn* conformation, and the proton wire Chr-Wat-Sr205-Glu222 affords efficient proton shuttling. In the B structures, the Glu222 side chain is in the *anti* conformation, thus disrupting the proton wire. In agreement with the results of spectroscopic measurements,[17−23] we found multiple conformations on the ground-state energy surface that may be referred to the I-type intermediates.

In sum, this work presents the first comprehensive quantum chemical characterization of the structural forms of the chromophore-containing domains of GFP in the ground electronic state that are consistent with all available experimental data.

ASSOCIATED CONTENT

Supporting Information

Coordinates of optimized structures. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

anem@lcc.chem.msu.ru; anemukhin@yahoo.com

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This study was partially supported by the Program on Molecular and Cell Biology from the Russian Academy of Sciences and by the Russian Foundation for Basic Research (project 13-03-00207). We acknowledge the use of supercomputer resources of the M.V. Lomonosov Moscow State University and of the Joint Supercomputer Center of the Russian Academy of Sciences. AIK acknowledges the support from the National Science Foundation via the CHE-0951634 grant. BLG thanks Dr. S.R. Belousov for valuable help.

REFERENCES

(1) Tsien, R. Y. *Annu. Rev. Biochem.* **1998**, *67*, 509−544.
(2) Zimmer, M. *Chem. Rev.* **2002**, *102*, 759−781.
(3) Remington, S. J. *Curr. Opin. Struct. Biol.* **2006**, *16*, 714−721.
(4) Craggs, T. D. *Chem. Soc. Rev.* **2009**, *38*, 2865−2875.
(5) Meech, S. R. *Chem. Soc. Rev.* **2009**, *38*, 2922−2934.
(6) Van Thor, J. J. *Chem. Soc. Rev.* **2009**, *38*, 2935−2950.

Table 2. Computed Wavelengths and Oscillator Strengths (in Brackets) of Electronic Transitions for the Major Isomers of the Structural Forms of wt-GFP and S65T Mutant

		S_0-S_1 (nm)			S_1-S_0 (nm)		
protein	structure	XMCQDPT2	SOS-CIS(D)	expt	XMCQDPT2	SOS-CIS(D)	exp
wt-GFP	neutral-wt-A1	388 [0.82]	390 [1.22]	395	461 [0.84]	455 [1.18]	440−460
	anion-wt-B1	484 [1.00]	489 [1.54]	475	519 [0.90]	536 [1.41]	
	anion-wt-I1	492 [0.97]	492 [1.45]	495	523 [0.92]	533 [1.40]	508
S65T	anion-S65T-B1	495 [1.03]	503 [1.51]	489	519 [0.97]	518 [1.47]	
	anion-S65T-I1	479 [0.96]	487 [1.43]		501 [1.01]	506 [1.44]	511

(7) Tonge, P. J.; Meech, S. R. *J. Photochem. Photobiol. A* **2009**, *205*, 1−11.
(8) Remington, S. J. *Protein Sci.* **2011**, *20*, 1509−1519.
(9) Chattoraj, M.; King, B. A.; Bublitz, G. U.; Boxer, S. G. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 8362−8367.
(10) Brejc, K.; Sixma, T. K.; Kitts, P. A.; Kain, S. R.; Tsien, R. Y.; Ormo, M.; Remington, S. J. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 2306−2311.
(11) Yang, F.; Moss, L. G.; Phillips, G. N. *Nat. Biotechnol.* **1996**, *14*, 1246−1251.
(12) Ormö, M.; Cubitt, A. B.; Kallio, K.; Gross, L. A.; Tsien, R. Y.; Remington, S. J. *Science* **1996**, *273*, 1392−1395.
(13) Palm, G. J.; Zdanov, A.; Gaitanaris, G. A.; Stauber, R.; Pavlakis, G. N.; Wlodawer, A. *Nat. Struct. Biol.* **1997**, *4*, 361−365.
(14) Jain, R. K.; Ranganathan, R. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 111−116.
(15) Shinobu, A.; Palm, G. J.; Schierbeek, A. J.; Agmon, N. *J. Am. Chem. Soc.* **2010**, *132*, 11093−11102.
(16) Creemers, T. M. H.; Lock, A. J.; Subramaniam, V.; Jovin, T. M.; Völker, S. *Nat. Struct. Biol.* **1999**, *6*, 557−560.
(17) Seebacher, C.; Deeg, F. W.; Bräuchle, C.; Wiehler, J.; Steipe, B. *J. Phys. Chem. B* **1999**, *103*, 7728−7732.
(18) Kennis, J. T.; Larsen, D. S.; van Stokkum, I. H.; Vengris, M.; van Thor, J. J.; van Grondelle, R. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 17988−17993.
(19) Van Thor, J. J.; Ronayne, K. L.; Towrie, M.; Sage, J. T. *Biophys. J.* **2008**, *95*, 1902−1912.
(20) Di Donato, M.; van Wilderen, L. J. G. W.; van Stokkum, I. H. M.; Stuart, T. C.; Kennis, J. T. M.; Hellingwerf, K. J.; van Grondelle, R.; Groot, M. L. *Phys. Chem. Chem. Phys.* **2011**, *13*, 16295−16305.
(21) Van Oort, B.; ter Veer, M. J. T.; Groot, M. L.; van Stokkum, I. H. M. *Phys. Chem. Chem. Phys.* **2012**, *14*, 8852−8858.
(22) Van Stokkum, I. H. M.; Gobets, B.; Gensch, T.; van Mourik, F.; Hellingwerf, K. J.; van Grondelle, R.; Kennis, J. T. M. *Photochem. Photobiol.* **2006**, *82*, 380−388.
(23) Fang, C.; Frontiera, R. R.; Tran, R.; Mathies, R. A. *Nature* **2009**, *462*, 200−205.
(24) Lill, M. A.; Helms, V. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 2778−2781.
(25) Warren, A.; Zimmer, M. *J. Mol. Graphics Model* **2001**, *19*, 297−303.
(26) Patnaik, S. S.; Trohalaki, S.; Pachter, R. *Biopolymers* **2004**, *75*, 441−452.
(27) Zhang, R.; Nguyen, M. T.; Ceulemans, A. *Chem. Phys. Lett.* **2005**, *404*, 250−256.
(28) Vendrell, O.; Gelabert, R.; Moreno, M.; Lluch, J. M. *J. Am. Chem. Soc.* **2006**, *128*, 3564−3574.
(29) Vendrell, O.; Gelabert, R.; Moreno, M.; Lluch, J. M. *J. Chem. Theory Comput.* **2008**, *4*, 1138−1150.
(30) Nemukhin, A. V.; Grigorenko, B. L.; Savitsky, A. P. *Acta Naturae* **2009**, *1*, 33−43.
(31) Bravaya, K. B.; Grigorenko, B. L.; Nemukhin, A. V.; Krylov, A. I. *Acc. Chem. Res.* **2012**, *45*, 265−75.
(32) Laino, T.; Nifosì, R.; Tozzini, V. *Chem. Phys.* **2004**, *298*, 17−28.
(33) Sinicropi, A.; Andruniow, T.; Ferre, N.; Basosi, R.; Olivucci, M. *J. Am. Chem. Soc.* **2005**, *127*, 11534−11535.
(34) Hasegawa, J.-Y.; Fujimoto, K.; Swerts, B.; Miyahara, T.; Nakatsuji, H. *J. Comput. Chem.* **2007**, *28*, 2443−2452.
(35) Virshup, A. M.; Punwong, C.; Pogorelov, T. V.; Lindquist, B. A.; Ko, C.; Martínez, T. J. *J. Phys. Chem. B* **2009**, *113*, 3280−3291.
(36) Bravaya, K. B.; Khrenova, M. G.; Grigorenko, B. L.; Nemukhin, A. V.; Krylov, A. I. *J. Phys. Chem. B* **2011**, *115*, 8296−8303.
(37) Filippi, C.; Buda, B.; Guidoni, L.; Sinicropi, A. *J. Chem. Theory Comput.* **2012**, *8*, 112−124.
(38) 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.
(39) Ding, L.; Chung, L. W.; Morokuma, K. *J. Phys. Chem. B* **2013**, *117*, 1075−1084.
(40) Warshel, A.; Levitt, M. *J. Mol. Biol.* **1976**, *103*, 227−249.
(41) Grigorenko, B. L.; Nemukhin, A. V.; Topol, I. A.; Burt, S. K. *J. Phys. Chem. A* **2002**, *106*, 10663−10672.
(42) Nemukhin, A. V.; Grigorenko, B. L.; Topol, I. A.; Burt, S. K. *J. Comput. Chem.* **2003**, *24*, 1410−1420.
(43) Gordon, M. S.; Freitag, M. A.; Bandyopadhyay, P.; Jensen, J. H.; Kairys, V.; Stevens, W. J. *J. Phys. Chem. A* **2001**, *105*, 293−307.
(44) Gordon, M. S.; Fedorov, D. G.; Pruitt, S. R.; Slipchenko, L. V. *Chem. Rev.* **2012**, *112*, 632−672.
(45) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. *J. Comput. Chem.* **1993**, *14*, 1347−1363.
(46) Gordon, M. S.; Schmidt, M. W. In *Theory and Applications of Computational Chemistry, the first forty years*; Dykstra, C. E., Frenking, G., Kim, K. S., ScuseriaG. E., Eds.; Elsevier: Amsterdam, 2005; pp 1167−1189.
(47) Ponder, J., http://dasher.wustl.edu/ffe/, accessed on February 16, 2013.
(48) Adamo, C.; Barone, V. *J. Chem. Phys.* **1999**, *110*, 6158−6170.
(49) Jacquemin, D.; Wathelet, V.; Perpète, E. A.; Adamo, C. *J. Chem. Theory Comput.* **2009**, *5*, 2420−2435.
(50) Granovsky, A. A. *J. Chem. Phys.* **2011**, *134*, 214113/1−14.
(51) Gozem, S.; Huntress, M.; Schapiro, I.; Lindh, R.; Granovsky, A.; Angeli, C.; Olivucci, M. *J. Chem. Theory Comput.* **2012**, *8*, 4069−4080.
(52) Granovsky A. A., FIREFLY Quantum Chemistry Package, version 7.1.G, http://classic.chem.msu.su/gran/firefly/index.html, accessed on February 16, 2013.
(53) Grimme, S. *J. Chem. Phys.* **2003**, *118*, 9095−9102.
(54) Rhee, Y. M.; Head-Gordon, M. *J. Phys. Chem. A* **2007**, *111*, 5314−5326.
(55) Shao, Y.; Molnar, L. F.; Jung, Y.; et al. *Phys. Chem. Chem. Phys.* **2006**, *8*, 3172−3191.
(56) Hsiao, Y.-W.; Sanchez-Garcia, E.; Doerr, M.; Thiel, W. *J. Phys. Chem. B* **2010**, *114*, 15413−15423.
(57) Hasegawa, J. Y.; Fujimoto, K. J.; Nakatsuji, H. *Chemphyschem* **2011**, *12*, 3106−3115.
(58) Stanton, J. F.; Gauss, J.; Ishikawa, N.; Head-Gordon, M. *J. Chem. Phys.* **1995**, *103*, 4160−4174.
(59) Melaccio, F.; Ferré, N.; Olivucci, M. *Phys. Chem. Chem. Phys.* **2012**, *14*, 12485−12495.