JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

Computational Study of Absorption Spectra of the Photoconvertible Fluorescent Protein EosFP in Different Protonation States

Petra Imhof*,†

Computational Molecular Biophysics, Interdisciplinary Center for Scientific Computing, University of Heidelberg, Speyererstrasse 6, 69115 Heidelberg, Germany

Supporting Information

ABSTRACT: Absorption spectra of the green-to-red convertible fluorescent protein EosFP have been computed in a hybrid quantum mechanical/molecular mechanical (QM/MM) framework. The experimentally observed absorption maximum at ∼390 nm is well reproduced by the protein with a neutral chromophore, and the anionic form is computed to absorb close to the experimentally determined maximum at ∼500 nm. Absorption of a zwitterionic form is calculated to lie in the same spectral region; however, this species cannot be unambiguously assigned to the experimental spectra. Variation of the protonation states of residues surrounding the chromophore do not have significant impact on the positions of the absorption maxima. In particular, protonation of Glu212 leaves the calculated spectra largely unaffected. This is consistent with the spectra of the E212Q mutant, which differ from the wild-type spectra only in the intensities but not in the positions of the absorption bands.

1. INTRODUCTION

Fluorescent proteins, the most popular of which is the green fluorescent protein (GFP), serve as labels for monitoring, for example, protein localization and protein interaction in vivo without being invasive.[1−3] GFP mutants with improved folding and photostability, higher brightness, and different colors have been engineered, and pH-sensitive photoactivatable and "kindling" variants have been developed; see, for example, ref 4 and references therein.

Green fluorescent protein and many of its relatives show different photophysical behavior at neutral and acidic pH. GFP, for example, is fluorescent at neutral pH, whereas at low pH (5−6), no light emission can be detected. The absorption spectra at lower pH show an additional band compared to the spectra taken at pH ∼ 7. This additional band at ca. 400 nm has been assigned to the neutral form of the chromophore, whereas the other band at ca. 470−480 nm has been attributed to an anionic form.[3,5−7]

Red fluorescent proteins have attracted particular interest because their emission wavelengths do not overlap with the green autofluorescence of living cells. Among the red fluorescing proteins, one can distinguish those which reach their red emission after a certain maturation time, like DsRed,[8,9] and those whose red emission is induced by illumination with blue or UV light, for example, Kaede and EosFP,[10−15] which leads to protein backbone cleavage, resulting in an extension of the chromophoric π-system. The efficiency of the conversion reaction is pH-dependent, suggesting that the protonation state of the protein and of the chromophore is crucial for formation of the photoconverted product.

The chromophore of the parent green fluorescent protein, most popularly modeled as 4-hydroxybenzylideneimidazolinone (HBI) or its 2,3-dimethyl variant (HBDI), has been subject of many theoretical studies, at various levels of theory, from semiempirical descriptions up to high-level, multireference calculations such as MRCI and CASPT2.[16−21] Computational approaches for describing the photophysical properties of the green fluorescent protein chromophore face a particular challenge, owing to the difficulty of describing several excited states and, at the same time, different protonated forms with similar accuracy. More recently, solvent effects or the protein environment have been included in computations on fluorescent proteins.[9,17,22−35]

Most studies focus on the neutral or anionic state of the chromophore, ruling out the protonated form. A zwitterionic species, in which an additional proton is located at the imidazole moiety of the chromophore while the hydroxyl group of the phenol moiety is deprotonated, has been considered in a few studies.[16,19,21,33] The zwitterionic chromophore has been reported to play a crucial role in the photoswitchable fluorescent protein asFP595.[33] On the basis of simulated absorption spectra of asFP595, supported by pK_a calculations, the trans wild type has a zwitterionic chromophore and the cis mutant of asFP595 has a neutral chromophore. A mechanism involving trans−cis isomerization of the chromophore, combined with several proton transfers, was proposed to underlie the photoactivation process.[33]

EosFP is a green-to-red photoconvertible fluorescent protein with a triad of residues (His62-Tyr63-Gly64) characterizing its chromophore. The photoactivated reaction leads to cleavage of the peptide bond between the amide group of residue 61 and the Cα atom of His62 and subsequent formation of a conjugated π-system, extending to His62 and giving rise to a red form of the protein. The structural basis for the reaction mechanism has been laid by Xray crystallography,[14] and quantum mechanical/molecular mechanical (QM/MM) computations have further elucidated the photoinduced cleavage mechanism.[36]

Received: August 9, 2012

This work focuses on the impact of different protonation states of the chromophore and different protonation patterns in the protein on the absorption spectra of the green form of EosFP and its E212Q mutant. Excitation energies of the isolated chromophore in neutral, anionic, and zwitterionic form have been computed at different levels of theory. Absorption spectra of the chromophore solvated in water and in the protein matrix are reported, resulting in an assignment consistent with previous suggestions.

2. METHODS

2.1. Electronic Excitation of the Chromophore in Vacuum. CASMP2/CASSCF(12,12) (CAS, complete active space; MP2, second-order Moller−Plesset perturbation theory; SCF, self-consistent field) vacuum calculations of the chromophore in different protonation states (neutral, anionic, and zwitterionic, cf. Figure 2) have been performed. These multiconfigurational self-consistent field (MCSCF) calculations have been carried out with the GAMESS-US program package.[37] The active space for optimization calculations comprises 12 electrons in 12 orbitals (12,12), leaving out the two occupied π-orbitals that are lowest in energy.

For the computation of electronic spectra in the protein environment, that is, computing the electronic transitions of an ensemble of conformations, MCSCF calculations are still too demanding. Two popular, computationally more efficient methods have been used: time-dependent density functional theory (TD-DFT) and a semiempiric, neglect of diatomic differential overlap (NDDO) -based method with multi-reference configuration interaction treatment (AM1/MRCI).

For the TD-DFT calculations, the program package Turbomole 5.9[38] has been applied, using the BP86 functional with resolution of the identity[39,40] and the SVP basis set.

The AM1/MRCI calculations have been carried out with a development version of the MNDO program.[41]

AM1/MRCI with three active spaces of different size has been applied: the full π-space with 16 electron in 14 orbitals (up to quadruple excitations, denoted here CISDTQ), and two smaller spaces, (12,12) and (8,8).

The computations of electronic excitation energies have been carried out as vertical excitations on the geometries optimized at the respective level of theory, except for the CASMP2 calculations, which have been carried out as single-point calculations on geometries optimized at the CASSCF level. Energies for adiabatic transitions have been calculated from optimized S_0-state and optimized S_1-state geometries. In addition, energies for de-excitation (emission) have been computed from the energy differences between excited state and ground state at the optimized S_1-state geometries.

Single-point calculations on geometries optimized on the RI-MP2/TZVP level of theory have been performed so as to allow a comparison of the description of the electronic transition independent of the quality of the geometry. RI-MP2 optimization calculations have been carried out with Turbomole. All geometries have been optimized to a gradient of $10^{-4}E_H$/bohr.

2.2. Absorption Spectra. Calculations of absorption spectra were performed from molecular dynamics (MD) ensembles. The molecular dynamics simulations have been carried out with the CHARMM force field by use of the programs CHARMM and NAMD[42] for simulations in a water shell and in the solvated protein, respectively. Parameters for the neutral and anionic chromophore were taken from ref 43. Parameters for the zwitterionic chromophore have been determined as follows: van der Waals parameters were the same as for the anionic and neutral forms. Atom charges have been derived from fitting water−chromophore hydrogen-bond lengths and interaction energies to those from Hartree−Fock/6-31G(d) calculations as suggested in ref 44. Water molecules have been placed at the two oxygen atoms as hydrogen donors and as hydrogen acceptor at the N2-bound hydrogen atom.

Bond and angle parameters were derived from fitting eigenvalues and eigenvectors from a normal-mode analysis to Hartree−Fock/6-31G(d) calculated normal modes by the procedure described in ref 45. Force constants from the neutral and anionic forms of the chromophore served as start parameters. Parameters for dihedral and improper angles are identical with those of the anionic chromophore. The final parameters are listed in the Supporting Information.

Molecular dynamics calculations were carried out for the chromophore and in the protein in a water shell of 20 Å radius. For simulations of the protein, only one subunit was considered, which was truncated at residue Asn221. As has been shown in ref 13, the spectroscopic properties of the monomeric protein resemble those of the tetrameric form.

The protein has been dissolved in a 90 Å × 90 Å × 90 Å large water box. For all molecular dynamics simulations, after 20 ps of heating and 300 ps of equilibration at 300 K by use of a Nose−Hoover thermostat, snapshots were taken every 1 ps from a 6 ns long simulation in case of the water sphere and every 2 ps from a 12 ns long simulation in the solvated protein; the first 1 and 2 ns, respectively, were omitted as equilibration phase.

On the snapshot structures, excitation energies were calculated by a hybrid quantum mechanical/molecular mechanical (QM/MM) approach, with TD-BP86 and AM1/MRCI applied for the QM part. The MM part was represented by the CHARMM force field. Three setups with differently sized quantum regions were considered. In the smallest setup, only the chromophore (Tyr64, Glu63 and His62 backbone) was treated quantum mechanically (green part in Figure 1), and the rest of the protein as well as the water molecules were explicitly taken into account as point charges. The water box

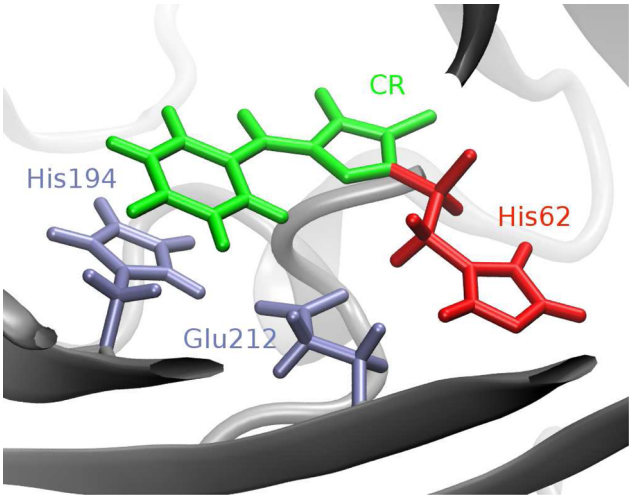

Figure 1. Residues of the fluorescent protein EosFP, treated quantum mechanically. Green, chromophore CR; red, additional inclusion of the His62 side chain; blue, QM region extended by side chains of His194 and Glu212. Gray cartoon indicates the protein environment.

dx.doi.org/10.1021/ct300706r | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

was truncated to a sphere consisting of the 4000 water molecules closest to the protein. For the intermediate setup, the QM region was extended to include also the side chain of the His62 residue (green + red part in Figure 1), and the largest setup included also the side chains of residues His194 and Glu212 in the quantum mechanically treated part of the system (green + red + blue part in Figure 1).

Absorption spectra were computed from the calculated excitation energies and oscillator strengths of the transitions to the first three electronically excited singlet states of the snapshots. With TD-DFT, the first five electronically excited singlet states were calculated. The spectra shown in Figures 3 and 4 are smoothed by representing each snapshot data point as a Gaussian curve with a width of 5 nm.

All calculations were carried out at the local group cluster, except for the molecular dynamics simulations, which were run on the Heidelberg Linux Cluster (Helics) of the Interdisciplinary Center for Scientific Computing.

3. RESULTS

3.1. Chromophore in Gas Phase. The $\pi-\pi^*$ ($S_0 \rightarrow S_1$) transition energies calculated with TDDFT, AM1/MRCI, and CASMP2/CASSCF of the neutral, anionic, and zwitterionic form of the green chromophore (cf. Figure 2) are summarized in Table 1. At the TD-DFT (BP86/SVP) level of theory, the transition to the first excited state is a $n-\pi^*$ (highest occupied to lowest unoccupied molecular orbital, HOMO−LUMO) transition with vanishing oscillator strength for the neutral and ionic forms of the chromophore. The (HOMO − 1)−LUMO transition is a $\pi-\pi^*$ transition and has a significant oscillator strength.

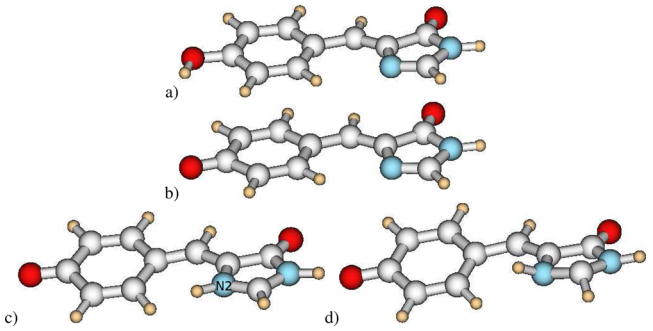

Figure 2. (a) Neutral, (b) anionic, and (c, d) zwitterionic forms of the chromophore of EosFP in (c) S_0-state and (d) S_1-state geometry.

Among the different methods applied, for the anionic state, the vertical excitation energies computed at the AM1/MRCI level of theory agree best with the experimental values. The deviation from experiment is only 0.04−0.12 eV in the case of the anionic chromophore, depending on the size of the active space. This is comparable to the accuracy of the CASMP2-(12,12) result (0.07 eV deviation). In the case of the neutral chromophore, the experimentally determined excitation energies have been measured for a model compound in which a positively charged spectator group (a methyleneammonium cation) was added to the neutral p-HBDI chromophore.[6] It has been shown[46] that addition of the charged group significantly shifts the absorption maximum to lower excitation energies. On the basis of high-level ab initio calculations, this shift is about 15 nm (0.15 eV).[46,47] Correction for this shift results in an estimated excitation energy of 355 nm (3.5 eV) for the neutral chromophore. The CASMP2 calculations overestimate this vertical excitation energy of the neutral form by 0.71 eV, whereas the AM1/MRCI calculations deviate by only 0.05−0.11 eV. Time-dependent DFT calculations, however, show better agreement with the experimental values for the neutral form than for the anionic form (0.24 and 0.45 eV deviation, respectively). For the zwitterionic form there is, to our knowledge, no experimental excitation energy available.

Adiabatic excitation energies are red-shifted to their vertical counterparts by 0.03−0.32 eV (neutral) and 0.07−0.2 eV (anionic) up to 0.16−0.46 eV (zwitterionic), the larger shift in the zwitterionic case being consistent with a more pronounced geometry change.

The optimized geometry of the neutral and anionic chromophores is almost planar in both the ground state and the electronically excited state. The zwitterionic chromophore, however, is slightly twisted in the ground state, which can be explained by the steric hindrance due to the extra hydrogen at

Table 1. Calculated Adiabatic and Vertical Excitation Energies of the First $\pi-\pi^*$ Transition ($S_0 \rightarrow S_1$) for Neutral, Anionic, and Zwitterionic Chromophores[a]

	neutral				anionic				zwitterionic			
	vertical		adiabatic		vertical		adiabatic		vertical		adiabatic	
	nm	eV	nm	eV	nm	eV	nm	eV	nm	eV	nm	eV
BP86/SVP[b]	380	3.26	399	3.12	408	3.04	418	2.97	450	2.75	499	2.48
AM1/MRCI(8,8)	343	3.61	366	3.38	503	2.47	521	2.38	557	2.22	600	2.06
AM1/MRCI(12,12)	359	3.45	363	3.42	486	2.55	500	2.48	541	2.29	673	1.83
AM1/MRCI(16,14, CISDTQ)	360	3.45	398	3.13	500	2.48	551	2.25	552	2.24	616	2.00
CASMP2(12,12)/ 6-31G(d,p)	294	4.21	318	3.90	492	2.52	533	2.32	425	2.92	501	2.55
experiment	370[c]	3.35[c]			479[d]	2.59[d]			n/a	n/a		

[a]Vertical excitation energies are calculated on S_0-state geometries optimized at the respective level of theory, except for the CASMP2 calculations, for which CASSCF-optimized geometries have been used. The adiabatic values are computed with geometries optimized at the S_0 and S_1 state at the respective level of theory. CASMP2 energy evaluations are carried out on CASSCF-optimized geometries. [b]The $\pi-\pi^*$ transition is $S_0 \rightarrow S_2$. [c]Reference 6. For these experiments, a positively charged methyleneammonium group was added to the neutral p-HBDI choromphore. [d]Reference 7.

dx.doi.org/10.1021/ct300706r | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

Table 2. Vertical Excitation Energies of the First $\pi-\pi^*$ Transition ($S_0 \rightarrow S_1$) for Neutral, Anionic, and Zwitterionic Chromophores[a]

	neutral		anionic		zwitterionic	
	nm	eV	nm	eV	nm	eV
BP86/SVP[b]	369	3.36	402	3.08	440	2.81
AM1/MRCI(8,8)	341	3.63	481	2.55	515	2.41
AM1/MRCI(12,12)	351	3.52	483	2.57	538	2.30
AM1/MRCI(16,14,CISDTQ)	352	3.52	496	2.50	536	2.31
CASMP2(12,12)/6-31G(d,p)	295	4.20	504	2.46	501	2.48
XMCQDPT2/cc-pVDZ//PBE0/cc-pVDZ[c]	375	3.31	494	2.51	503	2.46
sa-CASPT2(2,2)/6-31G(d)//B3LYP/6-31G(d)[d]	340	3.65	490	2.53	440	2.82
CASPT2(12,11)/cc-pVDZ//BLYP-cc-pVDZ[e]	345	3.58	465	2.67	n/a	n/a
experiment	370[f]	3.35[f]	479[g]	2.59[g]	n/a	n/a

[a]Calculated on RI-MP2/TZVP-optimized geometries, except if stated otherwise. High-level calculations of excitation energies from refs 21, 18, and 30 are given for comparison. HBDI denotes the dimethyl variant of the chromophore model used in this work. [b]The $\pi-\pi^*$ transition is $S_0 \rightarrow S_2$. [c]HBDI; ref 21. [d]Reference 18. [e] Reference 30. [f]Reference 6. For these experiments, a positively charged methyleneammonium group was added to the neutral p-HBDI chromophore. [g]Reference 7.

Table 3. Computed Energies for the Vertical $\pi^*-\pi$ Transition ($S_1 \rightarrow S_0$) for Neutral, Anionic, and Zwitterionic Chromophores[a]

	neutral		anionic		zwitterionic	
	nm	eV	nm	eV	nm	eV
BP86/SVP[b]	412	3.01	440	2.82	698	1.78
AM1/MRCI(8,8)	402	3.09	548	2.27	687	1.81
AM1/MRCI(12,12)	431	2.88	521	2.38	709	1.75
AM1/MRCI(16,14,CISDTQ)	465	2.67	558	2.22	633	1.96
CASMP2(12,12)/6-31G(d,p)	330	3.76	534	2.32	573	2.17
XMCQDPT2/SA-CASSCF(14,12)/cc-pVDZ//CASSCF(12,11)/cc-pVDZ[c]	459	2.70	523	2.37	n/a	n/a

[a]The energies are calculated on S_1-state geometries optimized at the respective level of theory, except for the CASMP2 calculations, for which CASSCF-optimized geometries have been used. [b]The $\pi^*-\pi$ transition is $S_2 \rightarrow S_0$. [c] HBDI; ref 21.

nitrogen atom N2. In the first electronically excited state, it shows a remarkable deviation from planarity; the N2 atom and the hydrogen atoms bound to N2 and the neighboring carbon atom C2 are significantly displaced out of the ring plane (see Figure 2). Table 3 lists the calculated energies for the vertical $\pi^*-\pi$ ($S_1 \rightarrow S_0$) transitions. All emission energies are significantly lower than those computed for absorption. The largest effect is a red shift of up to 1.8 eV, observed for the zwitterionic chromophore, which is consistent with deformation of the imidazole ring in the S_1 state.

Table 2 shows a comparison of excitation energies calculated at different levels of theory, computed on geometries that all have been optimized by use of RI-MP2/TVZ. The agreement of the computed values with the experimental ones is comparable to the excitation energies, which are calculated at geometries optimized at different levels of theory. Most excitation energies show a geometry dependence of ~0.1 eV. An interesting exception is the zwitterionic chromophore. Its CASMP2-calculated excitation energy is 0.44 eV lower when computed at the RI-MP2 level geometry than when CASSCF-(12,12)-optimized.

At all levels of theory applied, the calculated excitation energies show almost the same order for the three protonation states of the chromophore: The neutral form exhibits the highest excitation energy, followed by the anionic form, and the zwitterionic form absorbs at the lowest energy. The only exception is the highest level of theory, CASMP(12,12), on which the excitation energies calculated for the anionic and zwitterionic form are almost the same (2.46 and 2.48 eV, respectively),when RI-MP2-optimized geometries are used.

The method, among those applied here, that shows the best and most balanced agreement between computed and experimentally determined excitation energies is found to be AM1/MRCI. With an active space of (12,12), the deviation from experiment is 0.17 eV for the neutral chromophore and only 0.02 eV for the anionic chromophore. A smaller active space of (8,8) still agrees within 0.28 and 0.04 eV in the neutral and anionic cases, respectively.

3.2. Chromophore in Solution. In order to simulate absorption spectra at finite temperature in solution, molecular dynamics simulations of the chromophore in aqueous solution have been performed, and excitation energies have been computed from snapshots of these simulations. The dynamics in solution result in absorption maxima that are shifted to longer wavelengths compared to the gas-phase values. In the AM1/MRCI case, excitation energies are shifted by 10 nm for the anionic and zwitterionic forms to 358, 504, and 581 nm (3.47, 2.46, and 2.14 eV), respectively. When TD-BP86 is used on the same ensemble of conformations, the absorption maxima are computed at 387, 456, and 511 nm (3.21, 2.72, and 2.43 eV) for neutral, anionic, and zwitterionic forms, respectively. The respective spectra are shown in Figure 3.

dx.doi.org/10.1021/ct300706r | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

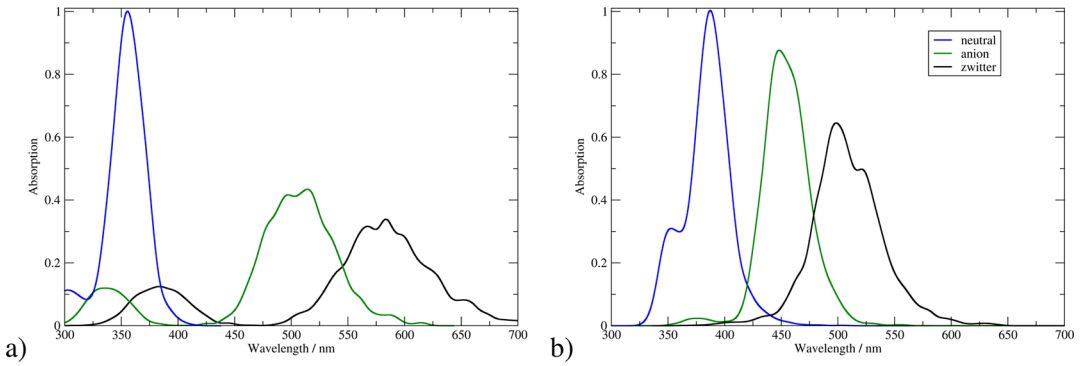

Figure 3. Simulated absorption spectra for the EosFP chromophore in water in its neutral (blue), anionic (green), and zwitterionic (black) form, calculated with (a) AM1/MRCI/MM and (b) TD-BP86/MM.

Table 4. Occupancies of Hydrogen Bonds between the Chromophore and Residues of the Surrounding Protein, Computed from the Simulations at Different Protonation States[a]

		occupancy, %									
		neutral				anionic				zwitterionic	
donor	acceptor		hsp	glup	hsp, glup		hsp	glup	hsp, glup		hsp
Arg66-NH2	CR-O2		36		42	18	45	88	91		14
Arg66-NH1	CR-O2		15	44		95	46			67	22
Arg91-NH2	CR-O2	98	74	95	71	88	83	86	80	95	85
CR-ND1	Gln38-OE1	10				42		30	64		
CR-NE2	Leu210-O		93		90		94		95		94
Ser142-OG	CR-OH		61	78	45	34	69	94	93		
CR-N2	Glu212-OE2							20		65	88
His194-ND1	Glu212-OE2	100	71	30	89	88	98	97	87	99	99

[a]CR, chromophore; hsp, His62 protonated; glup, Glu212 protonated.

3.3. Hydrogen-Bonded Environment of the Chromophore in the Protein. Table 4 lists the most stable hydrogen bonds formed between the chromophore and residues surrounding it. The carbonyl oxygen atom of the imidazole moiety, O2, forms a very stable hydrogen bond to Arg91 and in the anionic state also to Arg66. When protonated, a very stable hydrogen bond between the Nε atom of His62 and the backbone oxygen atom of the neighboring Leu208 is formed. Neutral His62 forms a hydrogen bond between the Nδ atom and Gln38 instead, if the chromophore is anionic. In all setups, except for the zwitterion, the oxygen atom of the chromophoric phenol moiety (former Tyr64) is hydrogen-bonded to Ser142.

In the crystal structure of EosFP, Glu212 is located below the imidazole moiety of the chromophore, in hydrogen-bonding distance to His194 (cf. Figure 1). The hydrogen bond to His192 is kept most of the time in all MD simulations, except for the setup with a neutral chromophore and protonated Glu212.

The zwitterionic chromophore forms a stable hydrogen bond to the negatively charged Glu212. The hydrogen bond between Glu212 and His194, however, remains formed for most of the simulation time, with Glu212 bridging between the chromophore and His194. The MD simulation of the anionic chromophore with protonated, neutral Glu212 shows 20% occupancy for an additional hydrogen bond between the N2 atom of the chromophore and Glu212. Although this is only a small fraction of the simulation time, it shows that Glu212 has a certain probability to deviate from the conformation observed in the crystal structure.

3.4. Absorption Spectra of the Protein. Absorption spectra of the protein have been computed for a number of different protonation states. In addition to variation of the protonation state of the chromophore (neutral, anionic, or zwitterionic), His62 has been simulated in either neutral or protonated form, and Glu212 has been simulated in anionic or neutral protonation state. Spectra for all protonation states have been computed with differently sized QM regions. Figure 4 shows absorption spectra computed with the large QM region. Depending on the size of the QM region, the computed absorption spectra show small differences. Their comparison is presented as Supporting Information.

The AM1/MRCI/MM-simulated spectra show absorption maxima at 369, 562, and 670 nm (3.36, 2.21, and 1.85 eV), corresponding to the electronic transition to the first excited singlet state in the neutral, anionic, and zwitterionic chromophore, respectively (see Figure4 a). The spectrum of the zwitterionic form shows a second band with a maximum at 450 nm (2.76 eV), corresponding to the $S_0 \rightarrow S_2$ transition . The spectra computed with TD-BP86/MM show electronic transitions to the second excited singlet state ($\pi-\pi^*$ transition); the first transition is a $n-\pi^*$ transition with vanishing oscillator strength. The absorption maxima are computed at 405, 451, and 503 nm (3.06, 2.75, and 2.47 eV) for neutral, anionic, and zwitterionic chromohpore, respectively (see Figure 4 b). Variation of the protonation state of the surrounding residues His62 and Glu212 leads to only small changes in the positions of the absorption maxima. The TD-BP86/MM-computed spectra, however, show a red shift of ~10

dx.doi.org/10.1021/ct300706r | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

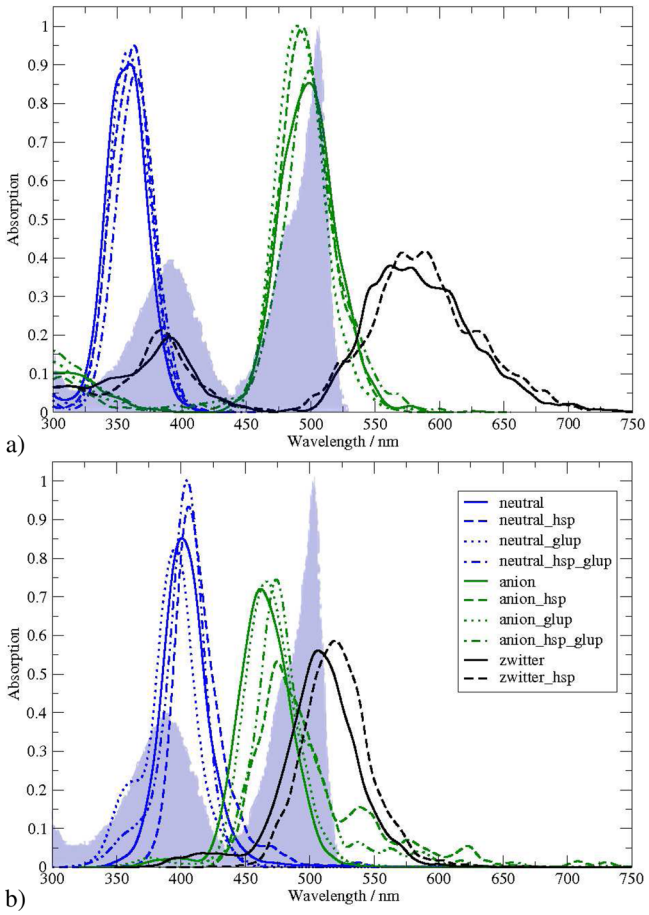

Figure 4. QM/MM-simulated absorption spectra of EosFP with the chromophore in neutral, anionic, and zwitterionic state with large QM region: chromophore + His62 + Glu212 + His192 computed at (a) AM1/MRCI/MM or (b) BP86/MM level of theory. Solid lines represent spectra from simulations with neutral His62, whereas dashed lines correspond to protonated His62 (hsp). Spectra from simulations in which Glu212 is protonated (i.e., neutral) are shown with dotted lines (glup) and those from simulations with an additional proton at His62 and at Glu212 (hsp_glup) are shown as dot-dashed curves. The gray shades indicate the experimental spectra, adapted from ref 13.

nm for protonated His62 in all three cases (neutral, anionic, and zwitterionic chromophore).

3.5. Absorption Spectra of the E212Q Mutant. In addition to spectra of the EosFP wild type, molecular dynamics calculations of the E212Q mutant of EosFP have been performed with neutral, anionic, and zwitterionic chromophores. Absorption spectra of the E212Q mutant, simulated from the ensembles with neutral chromophore, do not show any significant difference from those computed for the wild-type protein. The AM1/MRCI-calculated spectrum of the anionic chromophore is broader in the mutant and extends to longer wavelengths than the wild-type spectrum. The spectrum of the zwitterionic chromophore of the mutant is shifted to lower absorption energies by ∼50 nm compared to the wild-type spectrum.

4. DISCUSSION

4.1. Excitation Energies of the Chromophore. The order of the excited states calculated with TD-DFT is different compared to that from the multiconfiguration methods, AM1/MRCI and CASMP2; the lowest excited states result from a HOMO−LUMO $n-\pi^*$ transition with negligible oscillator strengths. The first bright transition is a (HOMO − 1)−LUMO $\pi-\pi^*$ transition to the second excited singlet state for all protonation states of the chromophore. Since, for the assignment of the spectra, only bright states are of interest, the false location of the dark state is only a minor drawback for the present work, and the TD-DFT results are included for comparison. On the other hand, the TD-DFT calculated excitation energies deviate from experiment in an unbalanced manner. This weakness can be observed for a variety of density functionals as reported in the thorough evaluation in ref 20.

The semiempirical AM1/MRCI method is observed to be the best-balanced method in terms of agreement of computed with experimental excitation energies. The sensitivity to the size of the active space is only weak and justifies the choice of the computationally most efficient AM/MRCI with eight electrons in eight orbitals for the computation of absorption spectra in the protein. In a recent study on the red fluorescent protein Dsred, a similar apporach (OM2/MRCI), although with a larger active space, has been reported to perform reasonably well compared to experimental values and computationally more demanding DFT/MRCI results.[34]

In numerous previous studies, excitation energies for the isolated chromophore have been reported, computed at various levels of theory.[16−18,21,30,34,46] However, the zwitterionic species has attracted less attention and we are only aware of only two studies reporting high-level calculations for the zwitterionic chromophore.[18,21] These two differ significantly in the values reported for the $S_0 \rightarrow S_1$ transition. To our knowledge, there is no experimentally determined excitation energy that could serve as a reference. According to ref 18, the zwitterionic species absorbs at a wavelength longer than the neutral and shorter than the anionic form (440 nm, 2.82 eV), whereas the excitation energies reported in ref 21 are lowest for the zwitterionic chromophore (503 nm, 2.46 eV), compared to neutral and anionic forms. In both cases, excitation eneries have been computed for DFT-optimized geometries (albeit with different functionals). In this work, excitation energies have been computed on geometries, optimized on RI-MP2 or CASSCF(12,12) level of theory. For the neutral and anionic chromophore, the resulting excitation energies are quite similar (within 0.1 eV). The excitation energies of the zwitterionic form, however, are rather sensitive to the underlying geometry. Optimization at RI-MP2 level yields a significantly lower excitation energy than the one computed with CASSCF. The first value is almost identical to the one reported in ref 21, and the other one is close to the energy reported in ref 18. Since both RI-MP2 and PBE0 can be assumed to perform better for the geometry optimization than CASSCF (the active space and hence the correlation treatment comprises only the π-orbitals and not the geometrically important σ-skeleton), the lower excitation energy values are considered to be more likely. Moreover, the AM1/MRCI calculations result in excitation energies close to the experimental values for the neutral and anionic chromophores. On this level of theory, the excitation energies of the zwitterionic chromophore are computed at 2.2−2.4 eV, depending on the choice of active space and geometry. This is again in good agreement with the excitation energy reported in ref 21.

4.2. Assignment of Protein Spectra. The impact of the size of the QM region in QM/MM treatments of GFP has been discussed in the literature; see, for example, refs 35 and 48. As also shown by the absorption spectra computed with smaller

QM regions (Supporting Information), anionic states are particularly sensitive to the treatment of their environment. The presently computed absorption spectra of EosFP hence face an intrinsic inaccuracy due to a limited QM treatment. Another possible source of errors is the use of a nonpolarizable force field for the MM part. As discussed in ref 35, polarization of the protein environment can account for 0.1−0.12 eV in excitation energies and leads to a significant increase of absorption intensities in the case of the neutral chromophore in GFP. However, an (additional) inaccuracy of 0.12 eV due to neglection of polarization effects is within the error range of the QM methods applied and should still allow for a distinction between the different protonation states of the chromophore.

In all setups applied here (different sizes of QM regions, different protonation states of surrounding residues), the absorption maximum of the neutral form is computed close to 390 nm. This is in agreement with earlier assignments, based on comparison to spectra of other fluorescent proteins (mainly GFP) and the fact that the band at 390 nm appears only in the experimental spectra taken at pH 5.5, and thus can unambiguously assigned to a neutral chromophore. The other band with a maximum at 506 nm and a shoulder, or second maximum, at 478 nm stems from another, ionic form of the chromophore that is present at both neutral pH and pH = 5.5.

According to the TD-BP86 calculations, both, anionic and zwitterionic forms have absorption maxima in that region (462 and 507 nm, respectively). A possible assignment involves two different protonation states of the chromophore, with the anionic and zwitterionic forms ascribed to the band shoulder at 478 nm and the band maximum at 506 nm, respectively.

According to the AM1/MRCI calculations, the most probable candidate for the ∼500 nm band is the protein with an anionic chromophore. The shoulder should then be attributed to a vibronic transition. This is in agreement with the assignment of absorption bands around 500 nm of other related fluorescent proteins.[3,5,28,49] Consequently, either the zwitterionic form is not present or, in contrast to the present calculations, it does not absorb in the respective spectral region.

Although no band in the experimental spectra can be clearly assigned to the zwitterionic form, its presence cannot be ruled out completely. The absorption of the zwitterionic form could be outside the recorded spectral region (i.e., at long wavelength beyond 600 nm) in the absorbance spectrum, or it could have an emission at much longer wavelengths than 700 nm, or a only weak signal that is covered by other bands, for example, that of the anionic form. According to our calculations, emission of the zwitterionic chromophore is significantly red-shifted with respect to its excitation wavelength. This red shift is consistent with the geometry change upon electronic excitation, which is significantly larger in the zwitterionic chromophore compared to the neutral and anionic forms. Furthermore, for the zwitterion, the oscillator strength of the $S_1 \rightarrow S_0$ transition is significantly reduced compared to the $S_0 \rightarrow S_1$ excitation, whereas the two other forms have comparable oscillator strengths for both transitions, excitation and emission. This suggests that the emission band of the zwitterionic species is shifted to longer wavelength, weakened, and most likely also considerably broadened and thus hardly detectable, especially if the zwitterionic state is only little populated.

A different assignment has been reported for the spectra of asFP595, where the band at ∼500 nm has been assigned to a zwitter-ionic form.[33] However, the zwitterionic form of asFP595 has a trans conformation, and the cis form is found to be much less stable, as opposed to EosFP. The spectra of the cis form of asFP595 show a band at ∼390 nm, which has been assigned to a neutral chromophore, similar to the assignment of the neutral form of EosFP. No 500 nm band was observed for the cis form of asFP595. This has been interpreted as the protons being either at the imidazole moiety and in solution (trans) or at the phenol site and at glutamate 215 (cis) but not at both parts of the chromophore, forming an unstable cationic chromophore or an anionic chromophore not protonated on either side.[33] The crystal structures of EosFP give no hint of a trans chromophore, and only the IrisFP variant has been observed to show significant cis−trans isomerization.[15]

4.3. Protonation States. For EosFP, one can consider a protonation equilibrium between the zwitterionic state and a state in which the chromophore is anionic and the proton is located at the amino acid Glu212. These two protonation sites are buried within the protein and are most likely only little affected directly by the pH of the solvent. However, the equilibrium between these two states is most likely influenced by the pH indirectly. At low pH, the phenol moiety of the chromophore is protonated and renders the protonation of a second site (the imidazole ring) to form a cationic chromophore unlikely. Vice versa, the probability of a zwitterionic chromophore to accept a (second) proton at the phenol moiety can be considered to be very low. Only when the proton is located at Glu212 is the probability of forming a neutral chromophore by accepting a proton at Tyr64 high enough, which might explain why the absorption band of the neutral form (at 390 nm) can be observed only at low pH = 5.5 .

The experimental spectrum of the E212Q mutant shows virtually no difference in excitation energies compared to the wild-type form,[50] which is reproduced by the computed spectra (Figure 5). The major difference is the band at 390 nm, which is present in the mutant form already at neutral pH. This has been explained by the pK_a value of Tyr64 being 0.8 pH unit higher in the mutant than in the wild-type protein.[50] In the E212Q mutant, the proton that is part of the amino group of Gln212 has a negligible probability for a transition to the imidazole moiety of the chromophore and hence the formation of a zwitterionic chromophore. This in turn renders protonation of the phenol moiety much more likely, consistent with the band of the neutral form present already at pH = 7.

The protonation states of residues surrounding the chromophore, His62 and Glu212, cannot be discriminated from the absorption spectra simulated with varying protonation states. The E212Q mutant does not show green-to-red photoconversion, suggesting a crucial role for Glu212 in the photoinduced reaction. According to the pathway calculations in ref 36, it is unlikely that Glu212 acts as a proton acceptor in the course of the cleavage reaction. However, a clear assignment whether Glu212 is anionic or protonated has not been made, nor has its role been reported. From comparison of computed with experimental spectra, the protonation state of Glu212 cannot be deduced either. According to the present calculations, both scenarios, unprotonated and neutral Glu212, lead to similar absorption spectra of the protein. The absorption spectrum of the E212Q mutant emphasizes that the position of the absorption maximum is insensitive to the charge of residue 212.

The species absorbing at the wavelength that subsequently leads to photoconversion clearly has a neutral chromophore. Whether His62 is neutral or protonated cannot be decided on

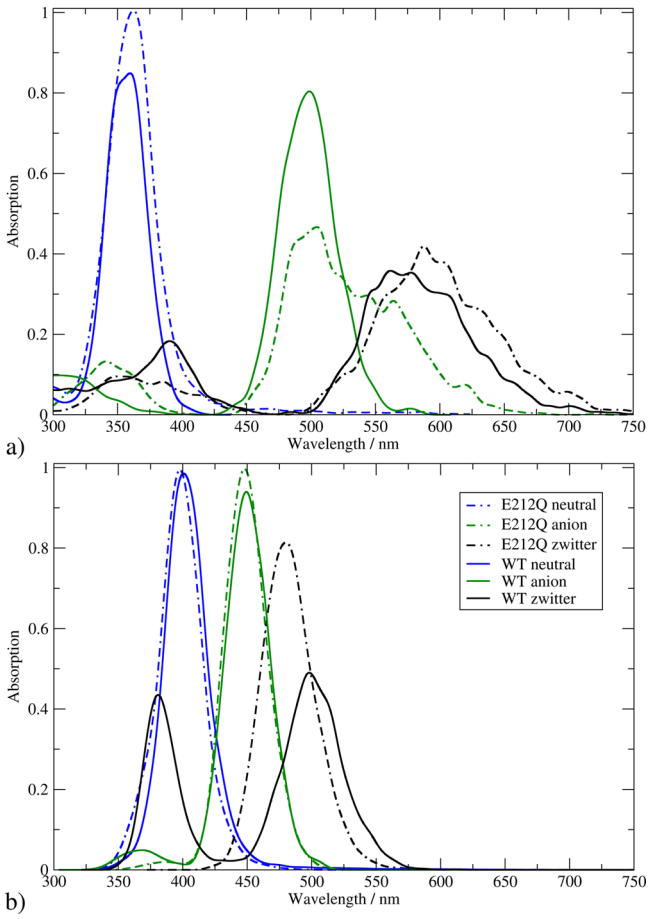

Figure 5. QM/MM-simulated absorption spectra of wild-type (WT, solid lines) and E212Q mutant (dash-dotted lines) EosFP with the chromophore in neutral (blue), anionic (green), and zwitterionic (black) states, with the large QM setup and (a) AM1/MRCI or (b) TD-BP86 as the QM method.

the basis of the computed spectra. The reaction pathway suggested in ref 36 involves a photoinduced proton transfer from His62 to Phe61 prior to peptide bond cleavage, which is putatively more likely with a protonated His62 in the reactant ground state.

The anionic chromophore does not play a role in the most likely photoconversion mechanisms suggested in ref 36, as opposed to earlier proposals.[14] It is, however, the species assigned to the dominating band in the absorption spectra of EosFP.

5. CONCLUSIONS

This paper reports absorption spectra computed for different protonation states of the chromophore and different protonation patterns in the green form of the photoconvertible fluorescent protein EosFP and its E212Q mutant. Variation of the protonation state of the surrounding residues His62 and Glu212 does not show a significant impact on the computed spectra. Mutation does not shift the band maxima, in agreement with experiment. According to the present calculations, the species absorbing at ∼390 nm has a neutral chromophore, in agreement with previous assignments.[13] The spectra of the protein with an anionic chromophore suggest an assignment of this species to the band at ∼500 nm. No band in the experimental spectra can be unambiguously assigned to the zwitterionic form. However, this species may participate in protonation equilibria between different states of the chromophore and its surrounding residues and thereby have an indirect impact on the intensity ratio of the experimentally observable bands. Clearly, further work, such as computing pK_a values, free energy differences, and transitions between differently protonated forms of the chromophore and its surrounding residues in EosFP, is necessary to elucidate the interplay of protonation equilibria between the various protonation states.

■ ASSOCIATED CONTENT

Supporting Information

Additional text, one figure, and seven tables giving absolute energies computed at the CASSCF and CASMP2 levels, occupation numbers of the active orbitals, force field parameters for the zwitterionic chromophore, and absorption spectra of the protein computed with differently sized quantum regions. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail petra.imhof@fu-berlin.de.

Present Address

[†]Theoretical Physics, Free University Berlin, Arnimallee 14, 14195 Berlin, Germany.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

I thank Frank Noé for his help with the Gaussian convolution of the computed spectra. Continuous support by Jeremy C. Smith is gratefully acknowledged.

■ REFERENCES

(1) Lippincott-Schwartz, J.; Snapp, E.; Kenworthy, A. *Nat. Rev. Mol. Cell Biol.* **2001**, *2*, 444−456.
(2) Prendergast, F. G. *Methods Cell Biol.* **1999**, *58*, 1−18.
(3) Zimmer, M. *Chem. Rev.* **2002**, *102*, 759−782.
(4) Subach, F. V.; Verkhusha, V. V. *Chem. Rev.* **2012**, *112*, 4308−4327.
(5) Sample, V.; Newman, R. H.; Zhang, J. *Chem. Soc. Rev.* **2009**, *38*, 2852−2865.
(6) Rajput, J.; Andersen, L. H.; Rocha-Rinza, T.; Christiansen, O.; Bravaya, K. B.; Erokhin, A. V.; Bochenkova, A. V.; Solntsev, K. M.; Dong, J.; Kowalik, J.; Tolbert, L. M.; Petersen, M. A.; Nielsen, M. *Phys. Chem. Chem. Phys.* **2009**, *11*, 9996−10002.
(7) Andersen, L. H.; Lapierre, A.; Nielsen, S. B.; Nielsen, I. B.; Pedersen, S. U.; Tomita, S. *Eur. Phys. J. D* **2002**, *20*, 597−600.
(8) Baird, G. S.; Zacharias, D. A.; Tsien, R. Y. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 11984−11989.
(9) Gross, L. A.; Baird, G. S.; Hoffmann, R. C.; Baldridge, K. K.; Tsien, R. Y. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 11990−11995.
(10) Anda, R.; Hama, H.; Yamamoto-Hino, M.; Mizuno, H.; Miyawaki, A. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 12651−12656.
(11) Dittrich, P.; Scha, S. P.; Schwille, P. *Biophys. J.* **2005**, *89*, 3446−3455.
(12) Mizuno, H.; Mal, T. K.; Tong, K. I.; Anda, R.; Furuta, T.; Ikura, M.; Miyawaki, A. *Mol. Cell* **2003**, *12*, 1051−1058.
(13) 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.
(14) Nienhaus, K.; Nienhaus, G. U.; Wiedenmann, J.; Nar, H. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 9156−9159.

(15) 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.
(16) Voityuk, A. A.; Michel-Beyerle, M.-E.; Rösch, N. *Chem. Phys. Lett.* **1997**, *272*, 162−167.
(17) Toniolo, A.; Olsen, S.; Manoha, L.; Martinez, T. *Faraday Discuss.* **2004**, *127*, 149−163.
(18) Olsen, S. The Electronic Excited States of Green Fluorescent Protein Chromophore Models. Ph.D. thesis, University of Illinois at Urbana−Champaign, Urbana, IL, 2004.
(19) Epifanovsky, E.; Polyakov, I.; Grigorenko, B. L.; Nemukhin, A. V.; Krylov, A. I. *J. Chem. Theory Comput.* **2009**, *5*, 1895.
(20) List, N. H.; Olsen, J. M.; Rocha-Rinza, T.; Christiansen, O.; Kongsted, J. *Int. J. Quantum Chem.* **2012**, *112*, 789−800.
(21) Polyakov, I. V.; Grigorenko, B. L.; Epifanovsky, E. M.; Krylov, A. I.; Nemukhin, A. V. *J. Chem. Theory Comput.* **2010**, *6*, 2377−2387.
(22) Sinicropi, A.; Andruniow, T.; Ferre, N.; Basosi, R.; Olivucci, M. *J. Am. Chem. Soc.* **2005**, *127*, 11534−11535.
(23) Helms, V. *Curr. Opin. Struct. Biol.* **2002**, *12*, 169−175.
(24) Lill, M. A.; Helms, V. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 2778−2781.
(25) Altoe, P.; Bernardi, F.; Garavelli, M.; Orlandi, G.; Negri, F. *J. Am. Chem. Soc.* **2005**, *127*, 3952−3963.
(26) Bravaya, K. B.; Khrenova, M. G.; Grigorenko, B. L.; Nemukhin, A. V.; Krylov, A. I. *J. Phys. Chem. B* **2011**, *115*, 8296−8303.
(27) Burgess, B. K.; Callis, P. R. *Biophys. J.* **1997**, *72*, A346 (abstract Th-AM-J3).
(28) Lopez, X.; Marques, M. A.; Castro, A.; Rubio, A. *J. Am. Chem. Soc.* **2005**, *127*, 12329−12337.
(29) Marques, M. A.; Lopez, X.; Varsano, D.; Castro, A.; Rubio, A. *Phys. Rev. Lett.* **2003**, *90*, No. 258101.
(30) Martin, M. E.; Negri, F.; Olivucci, M. *J. Am. Chem. Soc.* **2004**, *126*, 5452−5464.
(31) Nemukhin, A. V.; Grigorenko, B. L.; Savitskii, A. P. *Acta Nat.* **2009**, *2*, 41.
(32) Weber, W.; Helms, V.; McCammon, J. A.; Langhoff, P. W. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 6177−6182.
(33) Schäfer, L. V.; Groenhof, G.; Klingen, A. R.; Ullmann, G. M.; Boggio-Pasqua, M.; Robb, M. A.; Grubm üller, H. *Angew. Chem., Int. Ed.* **2007**, *119*, 536−542.
(34) Sanchez-Garcia, E.; Doerr, M.; Thiel, W. *J. Comput. Chem.* **2010**, *31*, 1602−1612.
(35) Steindal, A. H.; Olsen, J. M. H.; Ruud, K.; Frediani, L.; Kongsted, J. *Phys. Chem. Chem. Phys.* **2012**, *14*, 5440−5451.
(36) Lelimousin, M.; Adam, V.; Nienhaus, G. U.; Bourgeois, D.; Field, M. J. *J. Am. Chem. Soc.* **2009**, *131*, 16814−16823.
(37) 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., Jr. GAMESS, version 3 July 2003 (R2); Iowa State University, 2003.
(38) Ahlrichs, R.; Bar, M.; Haser, M.; Horn, H.; Kolmel, C. *Chem. Phys. Lett.* **1989**, *162*, 165−169.
(39) v. Arnim, M.; Ahlrichs, R. *J. Comput. Chem.* **1998**, *19*, 1746.
(40) Bauernschmitt, R.; Haser, M.; Treutler, O.; Ahlrichs, R. *Chem. Phys. Lett.* **1997**, *264*, 573.
(41) Thiel, W. MNDO, version 6.1; Max-Planck-Institut fuer Kohlenforschung, Mülheim an der Ruhr, Germany, 2004.
(42) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187−217.
(43) Reuter, N.; Lin, H.; Thiel, W. *J. Phys. Chem. B* **2002**, *106*, 6310−6321.
(44) Foloppe, N.; MacKerell, A. D. *J. Comput. Chem.* **2000**, *21*, 86−104.
(45) Vaiana, A. C.; Cournia, Z.; Costescu, I. B.; Smith, J. C. *Comput. Phys. Commun.* **2005**, *167*, 34−42.
(46) Filippi, C.; Zaccheddu, M.; Buda, F. *J. Chem. Theory Comput.* **2009**, *5*, 2074−2087.
(47) Lammich, L.; Petersen, M. A.; Nielsen, M. B.; Andersen, L. H. *Biophys. J.* **2007**, *92*, 201−207.
(48) Filippi, C.; Buda, F.; Guidoni, L.; Sinicropi, A. *J. Chem. Theory Comput.* **2012**, *8*, 112−124.
(49) Topol, I.; Collins, J.; Nemukhin, A. V. *Biophys. Chem.* **2010**, *149*, 78−82.
(50) Nienhaus, G. U.; Nienhaus, K.; Hölzle, A.; Ivanchenko, S.; Renzi, F.; Oswald, F.; Wolff, M.; Schmitt, F.; Röcker, C.; Vallone, B.; Heilker, W. W. R.; Nar, H.; Wiedenmann, J. *Photochem. Photobiol.* **2006**, *351*, 351−358.