JCTC Journal of Chemical Theory and Computation

Absorption Spectrum of the Green Fluorescent Protein Chromophore: A Difficult Case for ab Initio Methods?

Claudia Filippi,*[,†,‡] Maurizio Zaccheddu,[†] and Francesco Buda*[,§]

Instituut-Lorentz, Universiteit Leiden, Niels Bohrweg 2, 2333 CA Leiden, The Netherlands, Faculty of Science and Technology and MESA+ Research Institute, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands, and Leiden Institute of Chemistry, P.O. Box 9502, 2300 RA Leiden, The Netherlands

Received May 7, 2009

Abstract: We perform a thorough comparative investigation of the excitation energies of the anionic and neutral forms of the green fluorescent protein (GFP) chromophore in the gas phase using a variety of first-principle theoretical approaches commonly used to access excited state properties of photoactive molecules. These include time-dependent density functional theory (TDDFT), complete-active-space second-order perturbation theory (CASPT2), equation-of-motion coupled cluster (EOM-CC), and quantum Monte Carlo (QMC) methods. We find that all approaches give roughly the same vertical excitation for the anionic form, while TDDFT predicts an excitation for the neutral chromophore significantly lower than the highly correlated methods. Our findings support the picture emerging from the extrapolation of the Kamlet-Taft fit of absorption experimental data in solution and indicate that the protein gives rise to a considerable bathochromic shift with respect to vacuum. These results also open some questions on the interpretation of photodestruction spectroscopy experiments in the gas phase as well as on the accuracy of previous theoretical calculations in the more complex protein environment.

1. Introduction

Computational modeling is a crucial complement to experiments in deepening our understanding of optical processes occurring in photosensitive biological systems. Despite significant theoretical progress in electronic structure methods, it is far from trivial to accurately compute excitation properties of even relatively small photoactive molecules. It is not unusual in surveying the vast theoretical literature on photosensitive biosystems to find that the large spread of first-principle approaches used for a particular case yields an equally large spread of results and predictions.

The difficulties arise from the rather strict requirements the theoretical approach must meet to provide a predictive description of a photoactive biosystem. It should give an accurate and balanced description of the ground and the excited states of the photoactive site and also be able to treat a realistically large model of the biosystem. Time-dependent density functional theory (TDDFT)[1] is the most appealing approach to efficiently compute the excitations of large molecular complexes. However, known shortcomings of conventional adiabatic TDDFT as the underestimation of charge transfer excitations[2−4] or the lack of explicit inclusion of two- and higher-electron excitations[5−8] may severely affect the accuracy in the case of photoactive molecules. Therefore, despite their less favorable scaling with system size, highly correlated quantum chemical methods such as complete-active-space second-order perturbation theory (CASPT2)[9] have established themselves as the main theoretical tool to compute excitations of photosensitive systems[10,11] even for relatively large systems where compromises must be taken in the choice of the size of the active space or of the basis set.

The green fluorescent protein (GFP) first isolated in the jellyfish *Aequorea victoria*[12] is an intrinsically fluorescent

* Corrresponding author e-mail: filippi@lorentz.leidenuniv.nl (C.F.), f.buda@chem.leidenuniv.nl (F.B.).

† Instituut-Lorentz Leiden.

‡ Faculty of Science and Technology Twente.

§ Leiden Institute of Chemistry.

10.1021/ct900227j CCC: $40.75
Published on Web 07/24/2009

protein which has revolutionized cellular biology in the last decades.[13−15] Its most numerous and successful applications are as a genetic fusion partner to host proteins which maintain their normal functions but are then fluorescent and can be dynamically visualized in living cells and organisms. Beyond being extremely relevant in biotechnology, GFP represents a perfect playground for theoretical investigation of photoactive biomolecules due to several reasons. First, the spectroscopic features of the chromophore of GFP have been experimentally extensively studied in vacuo,[16−18] solution,[19−24] and in the protein environment,[25−30] offering a wealth of data for theoretical comparison. Then, GFP has already been the subject of a large number of computational semiempirical[31−33] and first-principle[34−46] studies, which often differ in the conclusions and do not therefore offer a consistent picture. For instance, early semiempirical calculations incorrectly assigned the charge state of the chromophore to the absorption bands in the protein.[31] On the other hand, first-principle TDDFT[42−44] and CASPT2[37] calculations significantly differ in the prediction of the absorption spectrum in vacuo, while these methods in combination with different quantum-mechanics/molecular-mechanics (QM/MM) schemes[36,38] yield results in agreement with each other and with the experimental spectrum in the more complex protein environment. These findings have opened new questions which require a more thorough comparative investigation already at the level of simple gas-phase models.

In this paper, we employ a variety of first-principle theoretical approaches to compute the electronic excitations in the gas phase of the neutral and anionic forms of the GFP chromophore. These are the protonation states relevant in the protein and responsible for the two room-temperature absorption peaks of wild-type GFP at 398 nm (3.12 eV) and 478 nm (2.59 eV), respectively.[29] In addition to the neutral and anionic chromophores, we consider a cationic model compound which was recently characterized in gas-phase spectroscopy experiments.[18] For all these models, we compare the performance of TDDFT and highly correlated techniques such as CASPT2, equation-of-motion coupled cluster (EOM-CC),[47] and the less widely used quantum Monte Carlo (QMC) methods.[48] We find that the treatment of the neutral form is more problematic for all theoretical approaches due to a stronger multiconfigurational nature of its electronic states as compared to the anionic case where a small complete-active-space (CAS) wave function is sufficient to converge both QMC and CASPT2 calculations. For the neutral chromophore, TDDFT predicts a vertical excitation significantly lower than the other correlated approaches, while all theoretical methods yield practically the same excitation energy for the anionic form. We note that a multistate approach has to be adopted in the CASPT2 calculations for the neutral chromophore and that our CASPT2 results for the anionic form are at variance with previous studies[37] using the same technique but a different definition of the zero-order Hamiltonian. Our findings are rather puzzling and raise concerns on the interpretation of photodestruction spectroscopy experiments in the gas phase[16−18] as well as on the accuracy of previous TDDFT[36] and CASPT2[38] calculations in the more complex protein environment. Differently from what had been inferred from these experiments and calculations, our theoretical excitations in the gas phase indicate that the bathochromic shifts in the protein are quite significant. Our results are consistent with values extrapolated from absorption experiments in solution.[22]

In Section 2, we briefly present the methods used in this paper and focus on the description of a novel QMC scheme where we simultaneously optimize orthogonal Jastrow-Slater wave functions for ground and excited states. Computational details are given in Section 3. The model chromophores are described in Section 4, and the results are presented in Section 5. Finally, discussion and conclusions are in Section 6.

2. Methods

In this paper, we employ a variety of first-principle quantum chemical methods. We omit here the description of the more traditional TDDFT, CASSCF, CASPT2, and CC approaches since these methods have become a well-established part of quantum chemistry, and their description can be found in appropriate textbooks.[1,49] In contrast, QMC methods[48] represent a less explored alternative to conventional highly correlated approaches and as such deserve a brief presentation. Moreover, the method used to optimize excited state QMC wave functions is novel and is therefore described in detail below.

2.1. QMC and State-Average Optimization of Multiple States. QMC methods provide an efficient and accurate description of both dynamical and static electronic correlation and have been mostly employed for ground-state studies of molecular and extended systems.[48] Their application to the computation of excited states of small photosensitive molecules is more recent and appears very encouraging.[50−53]

The key ingredient which determines the quality of a QMC calculation is the many-body trial wave function which, in the present work, is chosen of the Jastrow-Slater type. As we treat multiple states of the same symmetry, we write the ground- and excited-state wave functions as linear combination of spin-adapted configuration state functions (CSF) multiplied by a Jastrow correlation factor

$$\Psi^I = \sum_{i=1}^{N_{\mathrm{CSF}}} c_i^I \mathcal{J} C_i \tag{1}$$

where different states depend on their individual linear coefficients c_i^I but share a common set of single-particle orbitals and Jastrow factor $\mathcal{J}$. We use here a Jastrow factor which correlates pairs of electrons and each electron separately with a nucleus and employ different Jastrow factors to describe the correlation with different atom types. Since the optimal orbitals and expansion coefficients in Ψ^I may differ from the values obtained for instance in a CASSCF calculation in the absence of the Jastrow factor $\mathcal{J}$, it is important to reoptimize them in the presence of the Jastrow component.

Here, we present an efficient and simple approach to obtain accurate and orthogonal many-body wave functions for multiple states of the same symmetry.[52,53] We obtain the

Jastrow and orbital parameters which minimize the average energy over the state of interest and the lower states, while the linear coefficients in the CSF expansion ensure that orthogonality is preserved among the states. In the orbital and Jastrow optimization step, we extend the linear optimization method for ground states[54] to the optimization of multiple states in a state average fashion. The resulting scheme is simpler and superior to the approach we previously proposed[50] where only the orbitals were optimized and orthogonality was only approximately preserved among the states.

It is well-known that an optimal set of linear coefficients is readily obtained by solving the generalized eigenvalue problem

$$\sum_{j=1}^{N_{\mathrm{CSF}}} H_{ij} c_j^I = E_I \sum_{j=1}^{N_{\mathrm{CSF}}} S_{ij} c_j^I \tag{2}$$

The Hamiltonian and overlap matrix elements are estimated by a finite-sample average in variational Monte Carlo as

$$H_{ij} = \left\langle \frac{\mathscr{J}C_i}{\Psi_g} \frac{\mathscr{H}\mathscr{J}C_j}{\Psi_g} \right\rangle_{\Psi_g^2}, S_{ij} = \left\langle \frac{\mathscr{J}C_i}{\Psi_g} \frac{\mathscr{J}C_j}{\Psi_g} \right\rangle_{\Psi_g^2} \tag{3}$$

where the statistical average is over the Monte Carlo configurations sampled from Ψ_g^2. To obtain an accurate estimate of the optimal coefficients of more than one state, the guiding wave function Ψ_g should be chosen to have significant overlap with all states of interest. In particular, we use here $\sqrt{\sum_I |\Psi^I|^2}$. Importantly, it has been shown that the use of the nonsymmetric estimator of the Hamiltonian matrix of eq 3 yields a strong zero-variance principle and results in a particularly efficient approach.[55] The scheme has also the advantage to enforce orthogonality between the wave functions of the multiple states and ensure a generalized variational principle.

To optimize the nonlinear parameters in the Jastrow factor and orbitals of the multiple states, we follow a state-average (SA) approach to determine a set of orbitals and a Jastrow factor which give a comparably good description of the states under considerations while preserving orthogonality among the states. We alternate between optimizing the linear coefficients and the nonlinear (Jastrow and orbital) coefficients in which the quantity minimized is the weighted average of the energies of the states under consideration

$$E_{\mathrm{SA}} = \sum_{I \in A} w_I \frac{\langle \Psi^I | \mathscr{H} | \Psi^I \rangle}{\langle \Psi^I | \Psi^I \rangle} \tag{4}$$

where the weights w_I are fixed and $\sum_I w_I = 1$. Therefore, at convergence, the averaged energy E_{SA} is stationary with respect to all parameter variations subject to the orthogonality constraint, while the individual state energies E_I are stationary with respect to variations of the linear coefficients but not with respect to variations of the orbital or Jastrow parameters. In this approach, the wave functions are kept orthogonal and a generalized variational theorem applies.

To improve the orbital and Jastrow parameters at each SA iteration step, we extend the ground-state linear optimization approach[54] to the SA optimization of multiple states. Under a common variation in an orbital or Jastrow parameter p_i, the changes in the states Ψ^I are given by $\Psi_k^I = (\partial \Psi^I / \partial p_k)$ and can be made orthogonal to the corresponding state as

$$\Psi_k^I = \Psi_k^I - \left[\left\langle \frac{\Psi^I}{\Psi_g} \frac{\Psi_k^I}{\Psi_g} \right\rangle_{\Psi_g^2} \Big/ \left\langle \frac{\Psi^{I2}}{\Psi_g^2} \right\rangle_{\Psi_g^2} \right] \Psi^I \tag{5}$$

To linearize the minimization with respect to the nonlinear parameters, we work in the semiorthogonal basis of the functions $\{\bar{\Psi}_0^I, \bar{\Psi}_k^I\}$, where $\bar{\Psi}_0^I = \Psi^I$, and find the variations Δp_i in the parameters as the lowest eigenvalue solution of the generalized eigenvalue problem

$$H_{ij}^{\mathrm{SA}} \Delta p_j = E S_{ij}^{\mathrm{SA}} \Delta p_j \tag{6}$$

where $\Delta p_0 = 1$. The SA Hamiltonian matrix is computed as

$$H_{ij}^{\mathrm{SA}} = \sum_{I \in A} w_I \left[\left\langle \frac{\bar{\Psi}_i^I}{\Psi_g} \frac{H \bar{\Psi}_j^I}{\Psi_g} \right\rangle_{\Psi_g^2} \Big/ \left\langle \frac{\Psi^{I2}}{\Psi_g^2} \right\rangle_{\Psi_g^2} \right] \tag{7}$$

and an analogous definition holds for the SA overlap matrix elements. The matrix elements for all states are computed in a single variational Monte Carlo run with a guiding wave function Ψ_g chosen to have significant overlap with the states of interest. At convergence and for the optimal linear coefficients, the minimal energy E_{SA} (eq 4) is obtained: if the iterative scheme converges, the matrix elements H_{i0}^{SA} are zero, and, consequently, the derivatives of E_{SA} with respect to the parameter p_i are zero as they equal H_{i0}^{SA}.

In summary, one iteration of excited state optimization consists of the following steps: *i*) sample the quantities needed for the optimization of the linear coefficients with the appropriate guiding wave function; *ii*) diagonalize the matrix (eq 2) to obtain the optimal linear coefficients for the states under consideration; *iii*) sample for all states the quantities needed in the linear equations (eq 6) and obtain the parameters Δp_i; and *iv*) construct a set of improved orbitals and Jastrow parameters as $p_i \rightarrow p_i + \Delta p_i$. As in the optimization of a ground state wave function, when the nonlinear parameters are far from the optimal values, the optimization may need to be stabilized by shifting all diagonal elements except the first one as $H_{ij}^{\mathrm{SA}} \rightarrow H_{ij}^{\mathrm{SA}} + a_{\mathrm{diag}} \delta_{ij} (1 - \delta_{i0})$ where a_{diag} is automatically determined in a correlated sampling run as described in ref 54.

The trial wave functions are then used in diffusion Monte Carlo (DMC), which produces the best energy within the fixed-node approximation [i.e., the lowest-energy state with the same zeros (nodes) as the trial wave function].

3. Computational Details

The DFT calculations are carried out using the Gaussian 03[56] and the Amsterdam Density Functional (ADF)[57−59] code. The Gaussian code is used to optimize the ground state geometries of the chromophores within all-electron DFT with the BLYP[60] and B3LYP[61] functionals. The same code is also used to perform the all-electron linear-response adiabatic TDDFT calculations with the BLYP and B3LYP functionals at the corresponding ground state structures. All Gaussian calculations use a cc-pVTZ basis set and default convergence criteria. The ADF code is employed to perform all-electron

TDDFT calculations using the state-average orbital-dependent potential (SAOP)[62] in the eigenvalue difference matrix of the response equations, in combination with the LDA exchange-correlation potential in the linear response kernel and for the computation of the reference state. The Slater ET-pVQZ basis set is used in all ADF calculations.

The program MOLCAS 7.2[63] is used for the all-electron CASPT2 and multistate (MS) CASPT2[64] calculations. The starting SA-CASSCF wave functions are obtained with equal weights over the states of interest. We employ the default IPEA zero-order Hamiltonian,[65] and we indicate if an additional constant level shift[66] is added to the Hamiltonian. Most calculations are done with the cc-pVDZ basis set, but other basis sets such as 6-31G*, 6-31+G*, and cc-pVTZ are also tested. In the CASPT2 calculations, we do not correlate as many lowest orbitals of σ character as the number of heavy atoms in the model.

The EOM-CC calculations are performed with the code MOLPRO 2006.1.[67] The CC calculations include single and double excitations (CCSD) and employ both the cc-pVDZ basis and a basis we denote with cc-pVTZ′ where the cc-pVTZ basis is used for all atoms except hydrogen which is described by a cc-pVDZ basis. We do not correlate as many lowest orbitals of σ character as the number of heavy atoms in the model.

The program package CHAMP[68] is used for the QMC calculations. We employ scalar-relativistic energy-consistent Hartree−Fock pseudopotentials[69] where the carbon, nytrogen, and oxygen 1*s* electrons are replaced by a nonsingular *s*-nonlocal pseudopotential and the hydrogen potential is softened by removing the Coulomb divergence. We employ the cc-pVDZ Gaussian basis sets[69] constructed for these pseudopotentials. Different Jastrow factors are used to describe the correlation with a hydrogen, oxygen, and carbon atom. For each atom type, the Jastrow factor consists of an exponential of the sum of two-fifth-order polynomials of the electron−nuclear and the electron−electron distances, respectively.[70] The starting determinantal components are obtained in CASSCF calculations which are performed with the program GAMESS(US).[71] In all SA-CASSCF calculations, equal weights are employed for the states, and the final CAS expansions are expressed on the weighted-average CASSCF natural orbitals. The CAS wave functions of the ground and excited states may be truncated with an appropriate threshold on the CSF coefficients, and the union set of surviving CSFs for the states of interest is retained in the QMC calculations. The Jastrow correlation factor and the CI coefficients are optimized by energy minimization in a state-averaged sense within VMC as described in the previous Section, where equal weights are used in the optimization. When indicated in the text, also the orbitals are optimized along with the Jastrow and CI parameters. The pseudopotentials are treated beyond the locality approximation,[72] and imaginary time steps of 0.055 or 0.075 au are used in the DMC calculations.

4. Chromophore Models

The fluorescent chromophore of GFP consists of a *p*-hydroxybenzylideneimidazolinone[73] (*p*-HBI) which sits in a β-can structure formed by the protein[26] and, in wild-type GFP, exists in two protonation states, neutral and anionic, corresponding to a protonated (1C) and deprotonated (1A) phenolic oxygen, respectively.[19,25,27,30,32,35] The chromophore models studied in this paper are depicted in Figure 1 and can be divided in three groups: The anionic chromophores (A, B), the neutral chromophores (C, D), and a positively charged chromophore (E). We always refer to Figure 1 and its labels when describing the models in the rest of the paper.

The *p*-HBI anionic *minimal* model (A) is the smallest possible representation of the GFP chromophore and possesses C_s symmetry. Given its favorable size and symmetry, we study very extensively this model (and its neutral counterpart) using a variety of correlated techniques. This chromophore was also employed in previous correlated CASPT2 calculations,[37] but no experimental characterization is available. The anionic *methyl-terminated* model (B) has also C_s symmetry and only differs from the minimal model (A) in the termination with two methyl groups instead of the hydrogens in the imidazolone ring (*p*-HBDI). Even though we expect its electronic properties to be rather similar to those of model (A), we study this chromophore since it was synthesized and investigated in gas-phase photodestruction spectroscopy experiments, which place its absorption maximum at 2.59 eV,[16] that is, very close to the value of the anionic chromophore in the protein. The *p*-HBDI (B) model was also studied in absorption experiments in solution.[22]

For the neutral chromophore, we construct two models analogous to the anionic case, namely the minimal (C) and the methyl-terminated (D) model. To the best of our knowledge, no experiments are available for these neutral chromophores in the gas phase. The methyl-terminated (D) model was studied in absorption experiments in solution.[22]

The experimental photodestruction spectroscopy technique employed to study the anionic chromophore (B) makes use of an electrostatic ion storage ring and can therefore be applied only to charged molecules. Therefore, a positively charged compound (E) was synthesized by adding a methyleneammonium cation to the neutral *p*-HBDI choromphore (D) to mimic the spectrum of the neutral chromophore.[18] We refer to model (E) as the *neutral*$^+$ chromophore. In the photodestruction spectroscopy experiment, the absorption maximum of model (E) is located at 2.99 eV.

5. Results

5.1. Structural Analysis of the Models. The structures of the chromophores are relaxed in the ground state using DFT and various exchange-correlation functionals. We discuss in detail the geometrical features of the minimal anionic (A) and neutral (C) models as the same bond-length patterns are observed for the methyl-terminated models. The neutral$^+$ (E) chromophore requires a separate discussion. For the atom labeling, we refer the reader to Figure 2.

The bond lengths of the minimal anionic (A) and neutral (C) models optimized with the BLYP functional and a cc-pVTZ basis are shown in Figure 3. We only show the bond

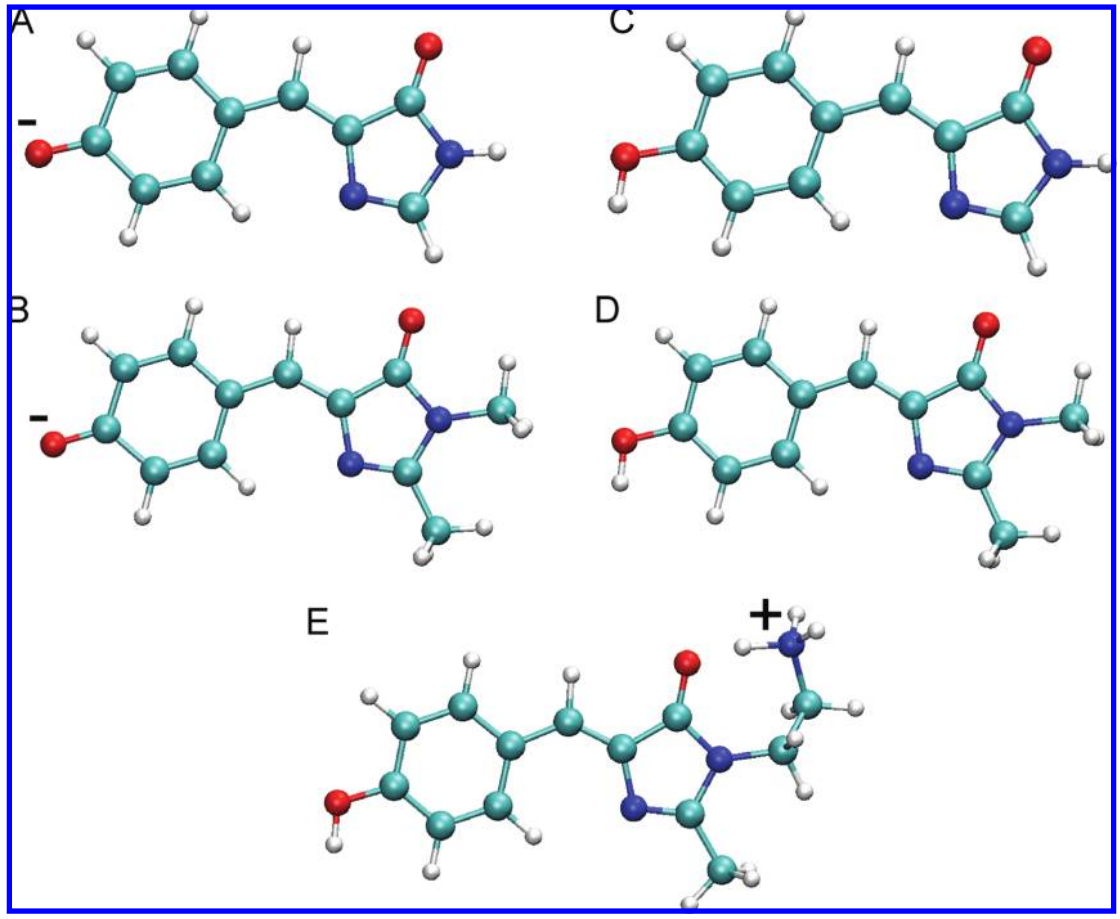

Figure 1. Gas-phase chromophore models: the anionic minimal (A) and methyl-terminated (B) models; the neutral minimal (C) and methyl-terminated (D) models; and the positively charged (neutral$^+$) model (E).

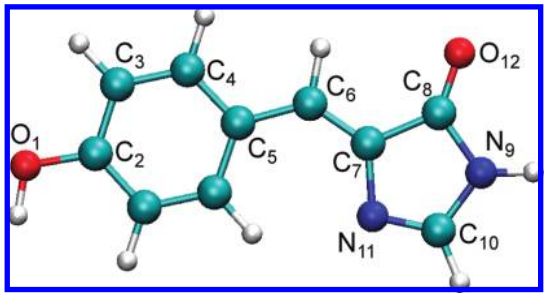

Figure 2. Atom numbering used for the chromophore of GFP.

lengths obtained with this exchange-correlation functional as other functionals yield practically equivalent geometries. In particular, the bond lengths obtained with the hybrid B3LYP functional are always shorter by only about 0.01 Å than the BLYP values.

Since the anionic chromophore may pose some difficulties for DFT, we also perform a CASPT2 optimization of the minimal anionic (A) model using a CAS(12,11) expansion (12 π-electrons distributed in 11 π-orbitals) and the 6-31G* basis set. The CASPT2 bond lengths are shown in Figure 3 where we also report the CASSCF bond lengths obtained in ref 37 using the same basis set and active space. We find that the CASSCF bond lengths are systematically shorter than the BLYP ones with an average deviation smaller than 0.02 Å and that the use of the CASPT2 method yields a geometry in better agreement with the BLYP one reducing the

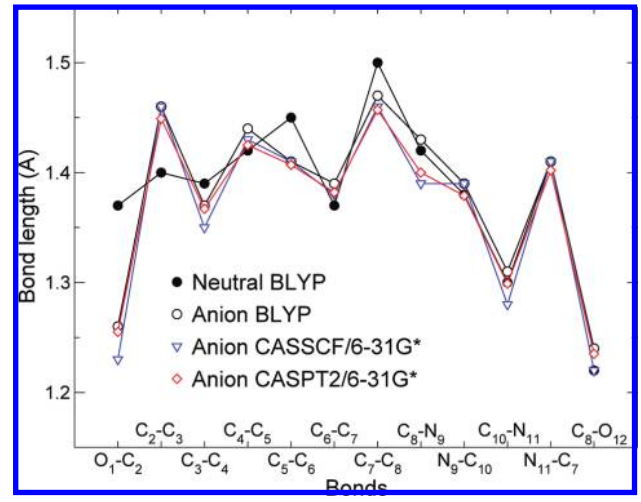

Figure 3. Bond lengths of the minimal anionic (A) and neutral (C) models optimized within DFT/cc-pVTZ with the BLYP functional. For the anionic model, the results from CASSCF/6-31G*[37] and CASPT2/6-31G* calculations are also shown. The bonds of the central carbon bridge are C_5-C_6 and C_6-C_7.

deviation to 0.01 Å. Therefore, since also the CASSCF and CASPT2 geometries are very close to the BLYP structure, we can safely claim that, when we compare excitation spectra for gas-phase models, the eventual differences one may observe should be attributed to the theoretical approaches

employed to compute the excitations rather then the approach adopted to optimize the geometry.

While the main structural features are largely independent of the theoretical method, the most evident difference is between the bond lengths of the anionic (A) and the neutral (C) model as shown in Figure 3. The geometrical changes between neutral and anionic chromophore can be rationalized in terms of the resonant benzenoid and quinoid forms which are accessible upon deprotonation of the phenolic oxygen.[33] The neutral model is characterized by the aromaticity of the phenolic ring with rather similar bond lengths and by a single−double bond-length alternation at the carbon bridge given by the bonds C_5-C_6 and C_6-C_7. In the anionic model, the phenolic oxygen is deprotonated, and the oxygen−carbon bond, O_1-C_2, shortens by about 0.1 Å as compared to the neutral model, losing its single-bond character. As a consequence, the aromaticity of the phenolic ring is reduced, yielding a quinoid-like structure of the ring. The degree of bond alternation in the central carbon bridge is also decreased in the anionic chromophore, where the two central bond lengths differ by only 0.02 Å as compared to a value of about 0.08 Å in the neutral chromophore. Beyond the central bridge, the deprotonation of the phenolic oxygen does not have a large effect.

The neutral$^+$ (E) chromophore represents a more complicated case due to the existence of two possible isomers. One isomer is depicted in Figure 1E and is characterized by the additional proton sitting on the NH_3^+ group, hydrogen bonded to the imidazolone oxygen. In the other isomer, the proton has jumped on the imidazolone oxygen leaving behind a neutral NH_2 moiety. In the paper where this compound was first studied, only the existence of the first isomer was postulated, and a minimal geometry was also obtained using MP3 techniques.[18] We find that, when optimizing the geometry of chromophore (E) with the BLYP functional, the structure with the proton on the NH_3^+ group is not stable as the proton prefers to bind to the oxygen. The use of the B3LYP functional yields instead two minima with the proton bound to the oxygen being energetically slightly favored by 2 mHartree. The BLYP and B3LYP geometries with the proton bound on the imidazolone oxygen are practically identical. Finally, we note that the bond length pattern of the neutral$^+$ chromophore has significant differences from both the anionic and the neutral models. Along the phenolic ring, the structure of the (E) model resembles the neutral form, but the carbon bridge is characterized by a much smaller bond-length alternation of only 0.03 Å, in line with the anionic value. Perhaps not surprisingly, the structure of the imidazolone ring of the (E) model differs from both the neutral and the anionic case due to the additional proton close to it.

5.2. TDDFT Excited States. We compute the low-lying singlet linear-response adiabatic TDDFT vertical excitations of the minimal (A, B), methyl-terminated (C, D), and neutral$^+$ (E) models. The BLYP and B3LYP functionals are employed on the corresponding BLYP and B3LYP geometries using the cc-pVTZ basis, while the SAOP excitations are computed only on the BLYP geometries with the ET-pVQZ basis.

Table 1. Adiabatic TDDFT Vertical Excitation Energies (eV) and Oscillator Strengths (in Parentheses) for the Minimal (A) and Methyl-Terminated (B) Anionic Models[a]

	BLYP	B3LYP	SAOP
Minimal Model (A)			
$S_0 \rightarrow S_1$	2.30(0.00) H-1→L(0.70)	3.16(0.00) H-1→L(0.69)	2.59(0.00) H-1→L(1.00)
$S_0 \rightarrow S_2$	**2.97**(0.75) H→L(0.54)	**3.16**(0.88) H→L(0.58)	**2.99**(0.79) H→L(0.95)
$-\varepsilon_{HOMO}$	0.59	1.29	5.10
Methyl-Terminated Model (B)			
$S_0 \rightarrow S_1$	2.29(0.00) H-1→L(0.70)	**3.09**(0.92) H→L(0.58)	2.58(0.00) H-1→L(1.00)
$S_0 \rightarrow S_2$	**2.89**(0.77) H→L(0.54) H→L+2(0.13) H→L+3(0.11)	3.16(0.00) H-1→L(0.69)	**2.93**(0.80) H→L(0.94)
$-\varepsilon_{HOMO}$	0.53	1.22	5.02

[a] The BLYP and B3LYP excitations are computed with a cc-pVTZ basis on the corresponding ground state geometries, while SAOP employs the BLYP geometries and the ET-pVQZ basis. The electronic transitions are listed if their weight (in parentheses) is greater than 0.1. The ionization threshold ($-\varepsilon_{HOMO}$) is also given. We use boldface characters to indicate the dominant excitation.

For all models, we report the TDDFT excitation energies with their oscillator strength and character. We also list minus the Kohn−Sham energy of the highest occupied molecular orbitals (HOMO) which gives the DFT ionization threshold. TDDFT may place a bound excitation in the continuum since the ionization threshold is underestimated in DFT due to the incorrect exponential decay of the Kohn−Sham potential for a pure functional.[74] The use of an asymptotically corrected exchange-correlation potential as SAOP cures this deficiency, raising the threshold to a significantly higher value, and allows us to verify that the excitation can in fact be trusted.[75]

In Table 1, we summarize the TDDFT results for the anionic minimal (A) and methyl-terminated (B) models. As expected, the electronic properties of these two chromophores are rather similar. The addition of the methyl groups only lowers the BLYP, B3LYP, and SAOP dominant excitations by 0.08, 0.07, and 0.06 eV, respectively, while preserving the same HOMO → LUMO character of the excitation and the same oscillator strength. The excitations obtained with B3LYP are consistently higher by about 0.2 eV than the BLYP values as can be expected for a hybrid functional, while SAOP gives excitations which are slightly higher than BLYP. For the methyl-terminated (B) model, the TDDFT excitation lies in the range 2.9−3.1 eV and is therefore significantly different from the location at 2.59 eV of the absorption maximum of photodestruction spectroscopy experiments.[16]

For both models, the dominant TDDFT excitation is a $\pi \rightarrow \pi^*$ transition and is not characterized by charge transfer as we can infer by inspecting the HOMO and LUMO orbitals in Figure 4. The HOMO orbital is π-bonding on both bonds of the central bridge and on the C_{10}-N_{11} bond of the imidazolone ring, while the LUMO is antibonding on these three bonds. As these orbitals are rather similar for the (A) and (B) models and all functionals, we only plot them for SAOP approach and the methyl-terminated (B) model.

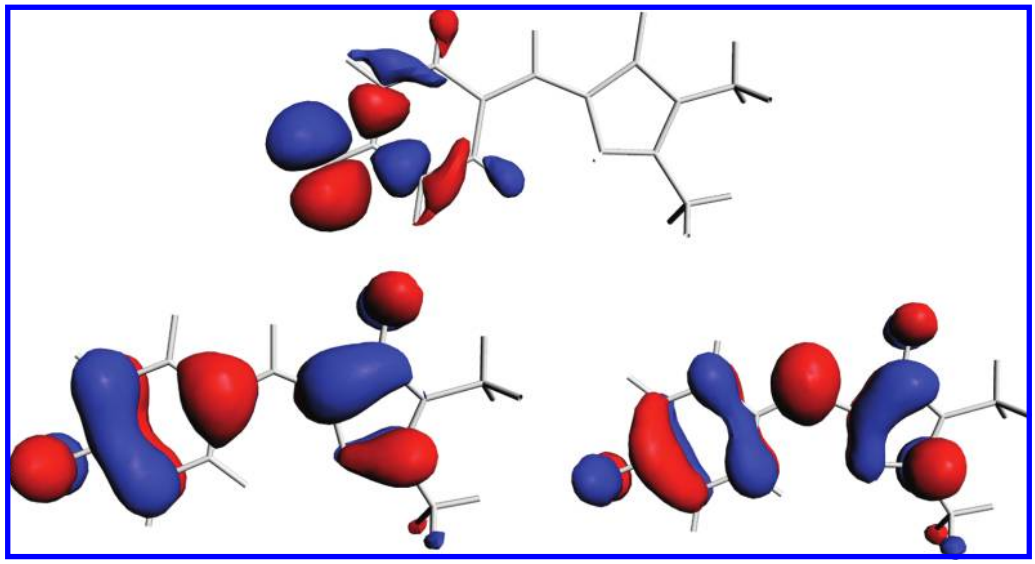

Figure 4. The DFT/SAOP orbitals for the anionic methyl-terminated (B) model. An isosurface of 0.03 is shown in red, and an isosurface of −0.03 is shown in blue. We only show the HOMO-1 (top), HOMO (bottom left), and LUMO (bottom right) orbitals, which are involved in the TDDFT/SAOP excitations.

Table 2. TDDFT Excitation Energies (eV) and Oscillator Strengths (in Parentheses) for the Minimal (C) and the Methyl-Terminated (D) Neutral Models[a]

	BLYP	B3LYP	SAOP
Minimal Model (C)			
$S_0 \rightarrow S_1$	2.88(0.00)	**3.54**(0.68)	3.13(0.00)
	H-1→L(0.70)	H→L(0.61)	H-1→L(1.00)
$S_0 \rightarrow S_2$	**3.22**(0.59)	3.56(0.00)	**3.28**(0.62)
	H→L(0.57)	H-1→L(0.69)	H→L(0.95)
	H-5→L(0.13)		
$-\varepsilon_{HOMO}$	4.98	5.85	9.50
Methyl-Terminated Model (D)			
$S_0 \rightarrow S_1$	2.99(0.00)	**3.46**(0.66)	**3.20**(0.61)
	H-1→L(0.70)	H→L(0.61)	H→L(0.94)
$S_0 \rightarrow S_2$	**3.10**(0.52)	3.66(0.00)	3.21(0.00)
	H→L(0.57)	H-1→L(0.69)	H-1→L(1.00)
	H-5→L(0.14)		
	H-2→L(0.13)		
$-\varepsilon_{HOMO}$	4.75	5.62	9.27

[a] See footnote *a* in Table 1 for further explanations.

We observe that the BLYP and B3LYP excitations of both anionic chromophores are significantly above the corresponding ionization threshold and lie in the continuum. The use of the SAOP potential cures the problem since the SAOP ionization threshold is well above the dominant excitation, which is now bound. Nevertheless, the TDDFT/SAOP excitation energies for the (A) and (B) models are very close to the BLYP values, and the picture remains practically unchanged with respect to the use of the BLYP functional.[75]

The TDDFT results for the neutral minimal (C) and methyl-terminated (D) models are summarized in Table 2. As in the anionic case, the electronic properties of the two chromophores are rather similar. The dominant excitation has HOMO → LUMO character and is only lowered by about 0.1 eV when the methyl groups are added. The B3LYP excitation energies are larger than the BLYP ones by more than 0.3 eV, while SAOP yields excitations which are at most 0.1 eV larger than the BLYP energies. The excitations of the neutral chromophores are larger than the excitations of the corresponding anionic models by roughly 0.2, 0.4, and 0.3 eV for the BLYP, B3LYP, and SAOP functionals, respectively.

Differently from the case of the anionic models, the underestimation of the DFT ionization threshold is not a concern for the neutral chromophores as this threshold is above the dominant excitation for both the BLYP and B3LYP functionals. The character of the excitations is $\pi \rightarrow \pi^*$ also in the neutral models as can be seen from the orbitals in Figure 5. We only plot the orbitals which are involved in the SAOP excitations for the neutral methyl-terminated (D) model since these orbitals are rather similar for the other model and the other functionals. The HOMO orbital has π-antibonding character on the first carbon bond of the central bridge, while it is π-bonding on the second bond. The situation is reversed in the LUMO where the sequence along the bridge is instead π-bonding/antibonding.

In Table 3, we report the TDDFT excitations for the neutral$^+$ (E) model. The B3LYP excitations for the two structural B3LYP minima with the proton on the imidazolone oxygen and with the proton on the NH_3^+ group only differ by 0.03 eV, so the two structures are energetically very close both in the ground and in the excited state. For the structure with the proton on the oxygen, BLYP yields an excitation energy lower by about 0.2 eV than the B3LYP value, while SAOP lies as always between BLYP and B3LYP. We note that, for this model, the DFT underestimation of the ionization threshold is not an issue. Photodestruction spectroscopy experiments place the absorption maximum at 2.99 eV,[18] reasonably close to the vertical BLYP excitation but 0.2 and 0.3 eV lower than the SAOP and B3LYP values, respectively.

5.3. CASPT2 Calculations. The CASPT2 vertical excitation energies are computed on the BLYP geometries of all models in Figure 1. Unless otherwise stated, the standard IPEA zero-order Hamiltonian[65] is used, and no additional constant level shift[66] is applied.

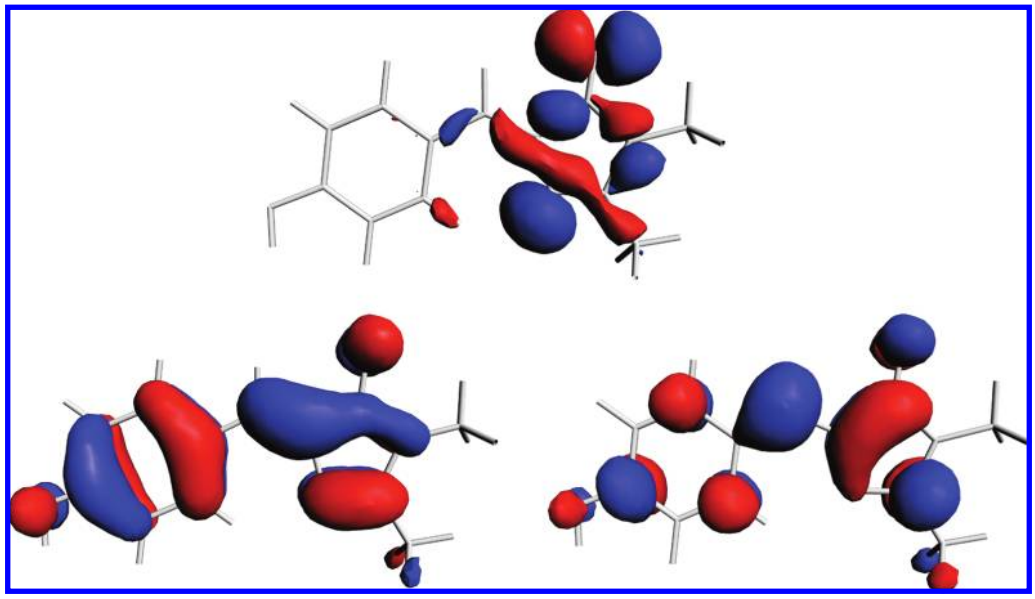

Figure 5. The DFT/SAOP orbitals for the neutral methyl-terminated (D) model. An isosurface of 0.03 is shown in red, and an isosurface of −0.03 is shown in blue. We only show the HOMO-1 (top), HOMO (bottom left), and LUMO (bottom right) orbitals, which are involved in the TDDFT/SAOP excitations.

Table 3. Adiabatic TDDFT Vertical Excitation Energies (eV) and Oscillator Strengths (in Parentheses) for the Neutral$^+$ (E) Model[a]

	BLYP	B3LYP	B3LYP	SAOP
	H^+ on O	H^+ on O	H^+ on N	H^+ on O
$S_0 \rightarrow S_1$	**3.07**(0.76)	**3.30**(0.92)	**3.33**(0.91)	**3.21**(0.69)
	H→L(0.55)	H→L(0.59)	H→L(0.60)	H→L(0.83)
	H-1→L(0.15)			H-1→L(0.15)
$-\varepsilon_{HOMO}$	8.29	9.17	8.74	12.79

[a] For the B3LYP functional, we list the results for the two geometries with the proton on the imidazolone oxygen (H^+ on O) and with the proton on the NH_3^+ group (H^+ on N). See footnote *a* in Table 1 and text for further explanations.

Table 4. CASPT2 Excitation Energies (eV) of the Anionic Minimal Model (A) Obtained Using Different Basis Sets and CAS(*n*, *m*) Expansions with *n* Electrons Distributed over *m* Orbitals of *a″* Symmetry (π Orbitals)[a]

CAS(*n*,*m*)	6-31G*	cc-pVDZ	cc-pVTZ
2,2	2.84	2.81	2.70
4,4	3.06	3.06	2.97
6,6	3.00	3.01	2.93
8,8	2.99	3.01	2.91
10,10	2.98	3.00	2.92
12,12	2.94	2.96	2.87
14,14	2.95	2.96	2.88
16,15	2.92	-	-

[a] The number of doubly occupied *a″* orbitals in the Hartree-Fock reference configuration is 8 for a total of 16 electrons. The DFT/BLYP ground state geometry is used.

The CASPT2 energies of the minimal anionic (A) model are listed in Table 4 for increasing size of the CAS expansion and different basis sets. The ground and the dominant excited state have A′ character, and the CAS(*n*, *m*) expansion to describe both states in a state-average CASSCF calculation is obtained by distributing *n* electrons over *m* orbitals of *a″* symmetry, that is, of π bonding/antibonding character. The system has 8 doubly occupied *a″* orbitals in the Hartree−Fock reference configuration, so 16 is the maximum number of π electrons which can be excited. We employ the cc-pVDZ and cc-pVTZ basis sets as well as the small 6-31G* basis set also used in the CASPT2/CASSCF(12,11) calculations on the same model in ref 37.

The dependence of the CASPT2 excitation energy on the size of the CAS is not very strong, and the result is already converged to better than 0.1 eV when a small CAS(6,6) expansion is employed. The character of the ground state is not strongly multiconfigurational, and the excited state is predominantly a HOMO → LUMO excitation. The oscillator strength computed using the state-average CASSCF wave functions is always greater than one. The 6-31G* and cc-pVDZ basis sets yield comparable excitations, while employing the cc-pVTZ basis systematically lowers the excitation by roughly 0.1 eV.

Our CASPT2 calculations therefore place the vertical excitation energy of the minimal anionic (A) model at about 2.9 eV. This result is at variance with previous CASPT2 calculations[37] which find an excitation energy of 2.67 eV employing the 6-31G* basis and a CASSCF(12,11) expansion. Our tests indicate that the basis and the size of the CAS do not influence the result to such an extent and that the discrepancy originates from the use of two different zero-order Hamiltonians. We employ the standard IPEA Hamiltonian implemented in MOLCAS 7.2, while the authors of ref 37 use a previous version of the code where the IPEA Hamiltonian was not available. If we perform a CASPT2 calculation with the CAS(14,14) expansion and set the IPEA shift to zero to reproduce similar conditions to ref 37, we find indeed an excitation of 2.63 eV.

For the methyl-terminated anionic (B) model, we find that CASPT2 predicts a very similar behavior as in the case of the minimal (A) model. The character and CASSCF oscillator strength of the electronic states are comparable, and, as a function of CAS size, the CASPT2 excitation energies are

Table 5. MS-CASPT2 Excitation Energies (eV) and CASSCF Oscillator Strengths (in Parentheses) of the Neutral Minimal (C) Model Obtained Using the cc-pVDZ Basis Set and CAS(*n,m*) Expansions with *n* Electrons Distributed over *m* Orbitals of *a″* Symmetry (π Orbitals)[a]

CAS(*n,m*)	energy (osc. strength)
10,10	4.13 (0.96)
12,12	4.03 (1.00)
14,13	3.77 (0.82)
14,14	3.78 (0.78)
16,15	3.72 (0.85)

[a] We use 4 states in the state-average CASSCF and MS-CASPT2 calculations, and the dominant state is always the 4th within CASSCF and the 2nd state within MS-CASPT2. We apply an additional level shift of 0.1 hartree in all CASPT2 calculations except in the CAS(16,15) calculation. The DFT/BLYP ground state geometry is used.

only slightly lower than the values of the minimal (A) model of Table 4. The CASPT2 excitation energy of the methyl-terminated anionic (B) model computed with a CAS(14,14) and a cc-pVDZ basis is equal to 2.92 eV. Photodestruction spectroscopy experiments locate the absorption maximum for this model at 2.59 eV.[16]

In Table 5, we show the CASPT2 excitation energies for the neutral minimal (C) model. The nature of the electronic states is different and more complex than in the anionic case since the dominant excited state is now strongly multiconfigurational and the CASPT2 result is very sensitive to the size of the CAS expansion. Moreover, the dominant CASSCF excited state is not always the lowest, but, as the size of the CAS increases, it becomes one of three excited states with relatively close CASSCF energies in a range of only 30 mHartree. Therefore, we need to perform multistate (MS) CASPT2 calculations over more than two states and to introduce an additional level shift to ensure convergence of the CASPT2 result. In the MS-CASPT2 calculations, the dominant excited state is the second one and is separated by more than 30 mHartree from the subsequent higher states, so the near degeneracy observed in CASSCF is lifted. In Table 5, we do not report the results for the smaller CAS expansions as the variations in the CASPT2 energy of the dominant state are as large as 0.3 eV as a function of the number of states included and the magnitude of the added level shift. Only with an expansion as large as a CAS(10,10) over 4 states and with a level shift of 0.1 hartree, we obtain a stable result in the sense that the dominant CASSCF state is the fourth one, an increase of the shift does not significantly affect the result, and employing 5 states with a shift of 0.2 hartree changes the excitation by only 0.02 eV. We note that this problematic convergence was ignored in a previous CASPT2 study of the neutral form using only a CAS(2,2) expansion over 2 states.[41] From Table 5, we observe that the convergence of the CASPT2 energy is not as smooth as in the anionic case, and it is necessary to correlate as many as 14 electrons to yield an apparently converged MS-CASPT2 result. As for the anionic case, we test the effect of setting the IPEA shift to zero and find that the MS-CASPT2 excitation with a CAS(14,14) expansion is lowered by as much as 0.45 eV when compared to the value given in Table 5.

Table 6. EOM-CCSD Vertical Excitation Energies in eV of the Anionic Minimal (A) and Methyl-Terminated (B), Neutral Minimal (C) and Methyl-Terminated (D), and Neutral$^+$ (E) Models[a]

model	cc-pVDZ	cc-pVTZ′
minimal anionic (A)	3.08	3.03
minimal neutral (C)	4.11	4.01
methyl-term. anionic (B)	3.04	-
methyl-term. neutral (D)	4.00	-
neutral$^+$ (E)	3.31	-

[a] We employ the cc-pVDZ and cc-pVTZ′ basis sets. The DFT/BLYP ground state geometries are used.

We also employ a CAS(14,14) expansion over 4 states to compute the excitation energy of the methyl-terminated neutral (D) model with the MS-CASPT2 approach and a level shift of 0.1 hartree. The resulting excitation of 3.58 eV is 0.2 eV lower than the CAS(14,14) value obtained for the minimal (C) model. Finally, we note that the excitation of the methyl-terminated neutral (D) model is significantly higher by as much as 0.7 eV than the excitation of the anionic (B) model.

Finally, we compute the CASPT2 excitation of the neutral$^+$ (E) model, using the BLYP geometry where the additional proton is bound to the oxygen of the imidazolone ring. It is not a priori obvious whether the behavior of the electronic states of the neutral$^+$ model are closer to the neutral or the anionic case since its geometrical features are distinct from either of them as discussed in Section 5.1. We find that the dependence of the CASPT2 excitation on the size of the CAS expansion is very smooth and similar to what we observe for the anionic model. The dominant excitation has strong HOMO $\rightarrow$ LUMO character, and a CAS(6,6) yields already a converged excitation to better than 0.1 eV. The CASPT2 excitation computed with a CAS(14,14) expansion and a cc-pVDZ basis is equal to 3.21 eV. Also in this case, setting the IPEA shift to zero reduces the excitation energy to 2.87 eV. We recall that photodestruction experiments locate the absorption maximum of the neutral$^+$ choromophore at 2.99 eV.[18]

5.4. CC Calculations. We compute the EOM-CCSD vertical excitation energies of all models in Figure 1. We employ the BLYP ground state geometries and the cc-pVDZ basis. For the (A) and (C) models, we also use the cc-pVTZ′ basis set.

The EOM-CCSD results are shown in Table 6. For all systems, the state with the largest oscillator stength is the second one and displays HOMO $\rightarrow$ LUMO character. We find that the use of a larger basis on the heavy atoms lowers the excitations of the minimal anionic (A) and neutral (C) models by only 0.05 and 0.1 eV, respectively. As expected, the excitation of the methyl-terminated anionic (B) model is very close to the value of the minimal anionic (A) model. The effect of deprotonation is instead very large as the difference between the excitation energy of the neutral (C) and anionic (A) chromophores is about 1.0 eV. The same large difference is observed between the excitations of the neutral and anionic methyl-terminated chromophores. Our EOM-CCSD results for the minimal models are in agreement with previous CC calculations.[41] Finally, the excitation

Table 7. Variational (VMC) and Diffusion Monte Carlo (DMC) Vertical Excitations in eV of the Minimal Anionic (A) Model, Computed with the cc-pVDZ Basis Set[a]

CAS(n,n)	Thr_{CSF}	N_{CSF}	VMC	DMC
2,2	0.00	3	3.41(4)	3.21(4)
4,4	0.00	20	3.39(4)	3.18(4)
6,6	0.10	7	3.18(4)	3.11(4)
8,8	0.10	9	3.14(4)	3.02(4)
8,8	0.05	25	3.18(4)	3.07(4)
8,8	0.02	75	3.21(4)	3.04(4)
10,10	0.05	27	3.20(4)	3.06(4)

[a] The statistical error on the last digit is indicated in brackets. We employ CAS(n,n) expansions obtained by distributing n electrons over n orbitals of a'' symmetry and expressed on the state-averaged CASSCF natural orbitals. We list the threshold applied to the CSF coefficients and the number of CSFs kept in the determinantal component of the QMC wave function. The CI and Jastrow parameters are optimized in a state-average minimization within VMC. The DFT/BLYP ground state geometry is used.

energy of the neutral$^+$ (E) is quite different than the value of the neutral model.

5.5. QMC Calculations. We compute the QMC vertical excitation energies of the minimal anionic (A), methyl-terminated (B), and neutral$^+$ (E) models. We employ the BLYP ground state geometries and the cc-pVDZ basis.

In Table 7, we present the VMC and DMC excitation energies of the minimal anionic (A) model, computed with different determinantal components in the trial wave function where only the CI coefficients and Jastrow parameters are optimized in a state-averaged minimization within VMC. As in the CASPT2 calculations, we compute the QMC excitations using CAS(n, n) expansions of increasing size obtained by distributing n electrons over n orbitals of a'' symmetry. We do not consider expansions larger than CAS(10, 10) since we have already demonstrated that the states of the anionic chromophore do not have a strong multiconfigurational character.

We notice that, starting with the CAS(6,6) expansion, the VMC excitation energy is reduced by about 0.2 eV with respect to the smaller CAS wave functions. The DMC excitation is also lowered by roughly the same amount when a CAS(8,8) is employed. In the CAS(8,8) calculations, we use three different thresholds to truncate the expansion and find equivalent VMC excitations in the range of 3.14−3.21 eV and in agreement with the CAS(6,6) value. The three corresponding DMC excitation energies are also compatible within one statistical error and fall in the range of 3.02−3.07 eV. Using a larger CAS(10,10) expansion truncated with a threshold of 0.05 yields a DMC excitation energy consistent with the CAS(8,8) results.

For the minimal anionic (A) model, we also investigate the effect of optimizing the orbitals together with the CI and Jastrow parameters and perform a full state-averaged optimization within VMC for the CAS(2,2) and CAS(4,4) wave functions as well as for the CAS(6,6) and CAS(8,8) wave functions truncated with a threshold of 0.1. Even though the gain in the VMC energy for each state can be as large as 30 mHartree, we find that, in all cases, the VMC and DMC excitation energies obtained with fully optimized QMC wave functions are compatible within one standard deviation with the corresponding values obtained by optimizing only the CI and Jastrow parameters.

Table 8. Vertical Excitations Energies (eV) of the Minimal Anionic (A) and Neutral (C), Methyl-Terminated Anionic (B) and Neutral (D), and Neutral$^+$ (E) Models[a]

model	DMC	SAOP	CASPT2	CCSD	PDS expt.	sol. expt.
anionic (A)	3.06(4)	2.99	2.96	3.08	-	-
anionic (B)	3.04(4)	2.93	2.92	3.04	2.59[b]	2.84[d]
neutral (C)	-	3.28	3.72	4.11	-	-
neutral (D)	-	3.20	3.58	4.00	-	3.51[d]
neutral$^+$ (E)	3.36(5)	3.21	3.21	3.31	2.99[c]	-

[a] We list the results computed with diffusion Monte Carlo (DMC), TDDFT/SAOP, CASPT2/cc-pVDZ, and EOM-CCSD/cc-pVDZ. The ground state DFT/BLYP geometries are used. We report the photodestruction spectroscopy (PDS) experimental absorption maxima from refs 16[b] and 18[c] as well as the results of the Kamlet-Taft multivariant fit of the absorption maxima in various solvents[22d].

For the methyl-terminated anionic (B) model, we compute the VMC and DMC excitation energies using wave functions obtained from a CAS(8,8) expansion truncated with a threshold of 0.05. We do not employ larger CAS expansions as we know that the CASSCF and CASPT2 calculations yield a well converged excitation energy already with a CAS(8,8) as discussed in Section 5.3. Moreover, the QMC calculations for the minimal (A) model indicate that the DMC excitations are well converged with this choice of CAS and threshold. The VMC and DMC excitation energies for the methyl-terminated anionic (B) model are equal to 3.21(4) and 3.04(4) eV, respectively.

Finally, we compute the QMC excitation energies of the neutral$^+$ (E) model using wave functions obtained from a CAS(8,8) expansion truncated with a threshold of 0.05 on the CSF coefficients. We find that the VMC and DMC excitation energies are equal to 3.43(5) and 3.36(5) eV, respectively.

6. Discussion and Conclusions

We will begin with the analysis of the relative performance of the various theoretical approaches and then compare with available experimental data. In particular, we will focus on the results from photodestruction spectroscopy experiments[16] which predict the bathochromic shift to be very small when the anionic chromophore is embedded in the protein as well as on the data in solution[22] where instead the absorption maximum of the neutral form is found to be significantly blue-shifted with respect to the protein.

To simplify the discussion, we summarize the key results in Table 8, where we report the vertical excitations of the minimal anionic (A) and neutral (C), the methyl-terminated anionic (B) and neutral (D), and the neutral$^+$ models. We recall that the minimal (A) and (C) models are sufficient to access the impact of deprotonation on the excitation energy, while, for the methyl-terminated (B) and neutral$^+$ (E) chromophores, photodestruction spectroscopy experiments are available.[16,18] For the methyl-terminated (B) and (D) models, we can also compare to absorption experiments in solution.[22] We list the TDDFT results obtained using the

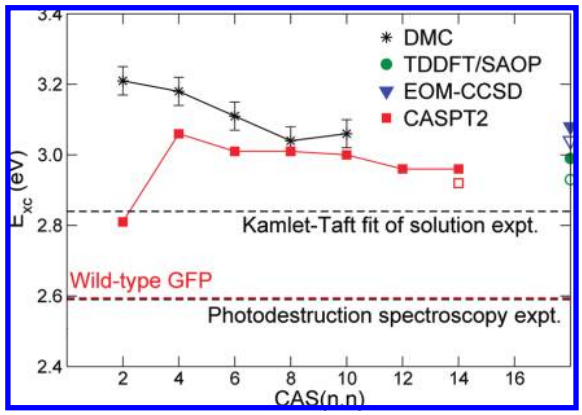

Figure 6. DMC and CASPT2 excitation energies (eV) of the anionic minimal (A) model obtained using increasing CAS(*n*, *n*) expansions. The EOM-CCSD and TDDFT/SAOP excitations are also reported on the right axis where the filled and open symbols refer to the anionic minimal (A) and the anionic methyl-terminated (B) model, respectively. The CASPT2 energy computed with the CAS(14,14) is also plotted for the (B) model using an open square symbol. The DFT/BLYP ground state geometry is used. For the methyl-terminated anionic (B) model, we indicate the location of the absorption maximum in photodestruction spectroscopy experiments[16] and the result of the Kamlet-Taft multivariant fit of the absorption maxima in various solvents.[22] The value of the absorption maximum in the wild-type GFP[29] is also plotted.

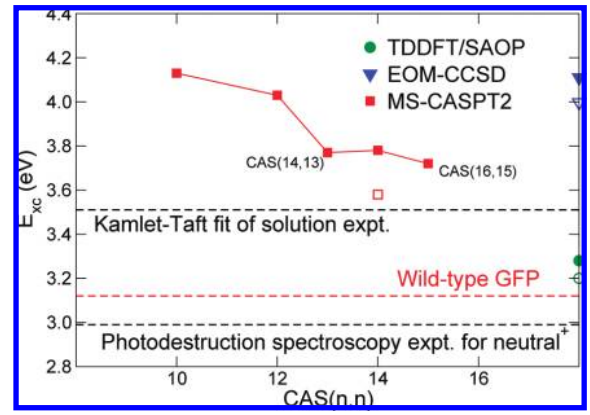

Figure 7. MS-CASPT2 excitation energies (eV) of the minimal neutral (C) model obtained using increasing CAS(*n*, *n*) expansions and 4 states. The EOM-CCSD and TDDFT/SAOP excitations are also reported on the right axis where the open symbols refer to the methyl-terminated (D) model. The MS-CASPT2 energy computed with the CAS(14,14) is also plotted for the (D) model using an open square symbol. For the methyl-terminated neutral (D) model, we indicate the result of the Kamlet-Taft multivariant fit of the absorption maxima in various solvents.[22] The value of the room-temperature absorption maximum in the wild-type GFP[29] is also plotted together with the location of the absorption peak of the neutral$^+$ (E) model measured in photodestruction spectroscopy experiments.[18]

SAOP approach since SAOP yields excitations which are between the BLYP and B3LYP results and cures the problem of the underestimation of the ionization threshold in the anionic models. For the CASPT2 and EOM-CC excitations, we only list the values obtained with the cc-pVDZ basis since results with this basis are consistently available for all models. We recall that the use of a larger basis set in these calculations is found to lower the excitation energies by roughly 0.05−0.1 eV. For the CASPT2 and DMC excitations, we report the values obtained with the largest CAS expansions.

We first observe that all theoretical approaches agree in predicting an excitation energy of the minimal anionic (A) model in the range of 2.96−3.08 eV. It is evident from Figure 6 that, when the quality of the wave function is improved, CASPT2 and DMC converge to a similar excitation, which is also in close agreement with the TDDFT/SAOP and EOM-CC values. We already noted that our CASPT2 excitation energy is significantly higher than the CASPT2 value of 2.67 eV reported in ref 37, and we attributed this discrepancy to the use of a different zero-order Hamiltonian. The IPEA Hamiltonian used in this work yields on average more accurate excitations, and its good performance appears to be here corroborated by the good agreement found between the CASPT2 and the DMC and EOM-CC excitations.

For all theoretical methods, the excitation energies of the methyl-terminated anionic (B) model are only slightly smaller than the corresponding values for the minimal anionic (A) model by about 0.05 eV. This is not surprising since the two chromophores only differ in the termination. We can therefore conclude that theory predicts an excitation of the methyl-terminated (B) model in the range of 2.92−3.04 eV. This estimate is rather robust since the effect of using a larger basis only affects the CASPT2 and EOM-CC excitations of the smaller minimal (A) model by less than 0.1 eV, and we may expect a similar behavior in the case of methyl terminations.

The minimal neutral (C) chromophore appears to be more difficult and controversial for theoretical methods than the anionic case. TDDFT/SAOP yields an excitation energy significantly smaller than the CASPT2 and EOM-CC values, which do also not agree with each other. TDDFT/SAOP places the vertical excitation of the neutral chromophore only 0.29 eV higher than the excitation of the anionic counterpart, while CASPT2 and EOM-CC predict a significant blue shift of 0.76 and 1.02 eV, respectively. The more complex nature of the excitation in the neutral chromophore is evident from Figure 7, where the convergence of the CASPT2 energy with the dimension of active space is clearly nontrivial: We employ the MS-CASPT2 approach over 4 states and do not even show the results for small CAS expansions given their dramatic dependence on the details of the calculations.

For the neutral$^+$ (E) model, all theoretical methods agree within 0.1 eV and predict an excitation energy in the range of 3.21−3.36 eV and, therefore, a similar shift of 0.27−0.32 eV between the excitations of the (E) and the anionic methyl-terminated (B) model. As it was the case for the anionic chromophore, the neutral$^+$ model does not pose any particular difficulty in the calculations. The states are not strongly multiconfigurational, and the CASPT2 and DMC excitations are well converged already when a small CAS expansion is employed. This must be contrasted to the situation encountered in the study of the neutral model. Therefore, even though the neutral$^+$ model was originally synthesized to

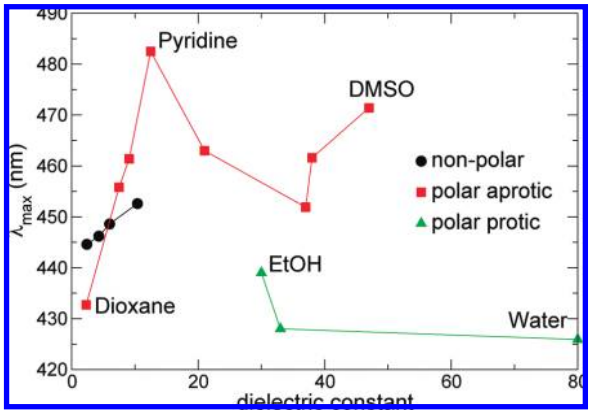

Figure 8. Experimental absorption maxima (nm) of the methyl-terminated anionic (B) model in solution from ref 22. The location of the peak is plotted as a function of the dielectric constant of the solvent. The solvents are classified as nonpolar, polar aprotic, and polar protic.

mimic the behavior of the neutral chromophore in photodestruction spectroscopy experiments,[18] its electronic states are very perturbed by the presence of the additional proton, and, consequently, this chromophore shares in fact very little with the neutral one.

To understand the link between our theoretical calculations and experiments, we report the relevant experimental data in Figures 6 and 7 as well as in Table 8. For the anionic chromophore, all theoretical approaches yield a vertical excitation energy in the gas phase significantly higher than the room-temperature absorption peak at 2.59 eV of the anionic form in the wild-type GFP protein.[29] Therefore, the predicted bathochromic shift is large, of the order of 0.3−0.4 eV, and the protein pocket must significantly modulate the excitation with respect to vacuum. This is in contrast to photodestruction spectroscopy experiments in the gas phase which predict an excitation very close to the value in the protein.[16] The disagreement between our threoretical results and photodestruction spectroscopy experiments is puzzling.

In an attempt to clarify this discrepancy, we also analyze recent extensive experimental data in solution[22] and plot the location of the absorption maximum of the anionic chromophore as a function of the dielectric constant of the solvent in 8. We note that the location of the absorption peak strongly depends on the solvent and spans a range as large as 60 nm (0.34 eV). While some solvents such as DMSO give very red-shifted excitations close to the value of 478 nm (2.59 eV) in the protein,[16] we can see at a glance that nonpolar solvents yield significantly blue-shifted values, and the extrapolation to the vacuum dielectric constant of one falls in the range 440 ± 5 nm (2.82 ± 0.03 eV). A more quantitative analysis is carried out in ref 22, where the multivariant Kamlet-Taft fit[76] of the absorption maxima in terms of the acidic, basic, and polar solvating parameters extrapolates to zero, that is, to the conditions of vacuum, at 2.84 eV. This value is only slightly lower than the vertical excitations computed with the various theoretical approaches, and this small red shift of 0.1−0.2 eV can easily be explained with a combination of basis set errors and vibronic effects which are very strong in this system. To estimate the energetic shift in the excited state due to vibronic effects, we perform a CASPT2 geometry optimization in the excited state, preserving C_s symmetry since the initial excited-state relaxation is believed to occur mainly in plane.[20,23] We find that the excitation is lowered by about 0.1 eV, a value compatible with a previous CASPT2/CASSCF study.[37]

For the neutral chromophore, the multivariant Kamlet-Taft fit of the absorption maxima in solution extrapolates to the conditions of vacuum, at 3.51 eV.[22] This value is significantly blue-shifted with respect to the absorption at 3.12 eV of the neutral form in the wild-type GFP protein, and no solvent yields an absorption sufficiently red-shifted to cover the value in the protein.[29] Our CASPT2 calculations support the picture emerging from the data in solution and yield an excitation of about 3.6 eV, slightly blue-shifted with respect to the extrapolation from the solution data but still perfectly compatible if basis set and vibronic effects are taken into account as in the case of the anionic chromophore. For the neutral case, there are no gas-phase experiments available, and we note that the absorption maximum observed in photodestruction spectroscopy experiments for the neutral$^+$ model[18] is rather far from the theoretical estimate obtained for the neutral chromophore and even slightly below the absorption value in the protein. We already pointed out that, theoretically, the neutral$^+$ model has a significantly lower excitation and little in common with the neutral chromophore in vacuum. Interestingly, we may expect this model to be closer to the situation in the protein where the Arg-96 positive counterion is in close proximity to the imidazolone oxygen. Therefore, the significant red-shift of the neutral$^+$ absorption in the direction of the protein value can be explained with the fact that neutral$^+$ does not mimic the neutral form in vacuum as originally suggested but possibly mimics the neutral chromophore in the protein environment.

In summary, we employ state-of-the-art theoretical approaches to investigate the absorption properties of several GFP chromophore models in the gas phase. Our theoretical findings nicely support the picture emerging from absorption experiments in solution and raise concerns on the interpretation of gas-phase photodestruction spectroscopy experiments especially for the anionic case. Our theoretical vertical excitations in the gas phase clearly indicate that the bathochromic shifts are quite significant and that the conditions of the chromphore in the protein are not close to those in vacuo. We finally note that our CASPT2 calculations differ from previous studies in the use of a more accurate zero-order Hamiltonian and that TDDFT differs significantly from highly correlated approaches in treating the neutral form of the chromophore.

Acknowledgment. We thank Mark Casida and Claudio Amovilli for many useful discussions. M.Z. is supported by the Stichting voor Fundamenteel Onderzoek der Materie (FOM). We acknowledge support from the Stichting Nationale Computerfaciliteiten (NCF-NWO) for the use of the SARA supercomputer facilities.

Supporting Information Available: DFT geometries of the model chromophores and CASPT2 geometry of the anionic minimal (A) model. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) *Time-dependent Density Functional Theory, Lecture Notes of Physics*; Marques, M. A. L., Ullrich, C. A., Nogueira, F., Rubio, A., Burke, K., Gross, E. K. U., Eds.; Springer: Berlin, 2006.

(2) Tozer, D. J.; Amos, R. D.; Handy, N. C.; Roos, B. O.; Serrano-Andrés, L. *Mol. Phys.* **1999**, *97*, 859–868.

(3) Casida, M. E.; Gutierrez, F.; Guan, J.; Gadea, F.-X.; Salahub, D. R.; Daudey, J.-P. *J. Chem. Phys.* **2000**, *113*, 7062–7071.

(4) Dreuw, A.; Weisman, J. L.; Head-Gordon, M. *J. Chem. Phys.* **2003**, *119*, 2943–2946.

(5) Casida, M. E. In *Recent Advances in Density Functional Methods, Part I*; Chong D. P., Ed.; World Scientific: Singapore, 1995; pp 155−188.

(6) Maitra, N. T.; Zhang, F.; Cave, F. J.; Burke, K. *J. Chem. Phys.* **2004**, *120*, 5932–5937.

(7) Cave, R. J.; Zhang, F.; Maitra, N. T; Burke, K. *Chem. Phys. Lett.* **2004**, *389*, 39–42.

(8) Casida, M. E. *J. Chem. Phys.* **2005**, *122*, 054111−1-8.

(9) Roos, B. O.; Andersson, K.; Fülscher, M. P.; Malmqvist, P.-A.; Serrano-Andrés, L. In *Advances in Chemical Physics*; Prigogine, I., Rice, S. A., Eds.; Wiley & Sons, Inc.: New York, 1996; Vol. 93, pp 219−331.

(10) Strambi, A.; Coto, P. B.; Frutos, L. M.; Ferré, N.; Olivucci, M. *J. Am. Chem. Soc.* **2008**, *130*, 3382–3388.

(11) Sinicropi, A.; Martin, E.; Ryazantsev, M.; Helbing, J.; Briand, J.; Sharma, D.; Léonard, J.; Haake, S.; Cannizzo, A.; Chergui, M.; Zanirato, V.; Fusi, S.; Santoro, F.; Basosi, R.; Ferré, N.; Olivucci, M. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 17642–17647.

(12) Shimomura, O.; Johnson, F. H.; Saiga, Y. *J. Cell. Comp. Physiol.* **1962**, *59*, 223–239.

(13) Tsien, R. Y. *Annu. Rev. Biochem.* **1998**, *67*, 509–544.

(14) Zimmer, M. *Chem. Rev.* **2002**, *102*, 759–781.

(15) Tozzini, V.; Pellegrini, V.; Beltram, F. In *Handbook of organic photochemistry and photobiology*; Horsphool, W. M., Lenci, F., Eds.; CRC Press: Washington, DC, 2004; Chapter *139*, pp 1−23.

(16) Nielsen, S. B.; Lapierre, A.; Andersen, J. U; Pedersen, U. V.; Tomita, S.; Andersen, L. H., *Phys. Rev. Lett.* **2001**, *87*, 228102−1-4.

(17) Andersen, L. H.; Bluhme, H.; Boyé, S.; Jorgensen, T. J. D.; Krogh, H.; Nielsen, I. B.; Nielsen, S. B.; Svendsen, A. *Phys. Chem. Chem. Phys.* **2004**, *6*, 2617–2627.

(18) Lammich, L.; Pertersen, M. A.; Nielsen, M. B.; Andersen, L. H. *Biophys. J.* **2007**, *92*, 201–207.

(19) Niwa, H.; Inouye, S.; Hirano, T.; Matsuno, T.; Kojima, S.; Kubota, M.; Ohashi, M.; Tsuij, F. I. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 13617–13622.

(20) Mandal, D.; Tahara, T.; Webber, N. M.; Meech, S. R. *Chem. Phys. Lett.* **2002**, *358*, 495–501.

(21) Vengris, M.; van Stokkum, I. H. M.; He, X.; Bell, A. F.; Tonge, P. J.; van Grondelle, R.; Larsen, D. S. *J. Phys. Chem. A* **2004**, *108*, 4587–4598.

(22) Dong, J.; Solntsev, K. M.; Tolbert, L. M. *J. Am. Chem. Soc.* **2006**, *128*, 12038–12039.

(23) Stavrov, S. S.; Solntsev, K. M.; Tolbert, L. M.; Huppert, D. *J. Am. Chem. Soc.* **2006**, *128*, 1540–1546.

(24) Webber, N. M.; Meech, S. R. *Photochem. Photobiol. Sci.* **2007**, *6*, 976–981.

(25) Chattoray, M.; King, B. A.; Bublitz, G. U.; Boxer, S. G. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 8362–8367.

(26) Ormö, M.; Cubitt, A. B.; Kallio, K.; Gross, L. A.; Tsien, R. Y.; Remington, S. J. *Science* **1996**, *273*, 1392–1395.

(27) Brejc, K.; Sixma, T. K.; Kitts, P. A.; Kain, S. R.; Tsien, R. Y.; Ormö, M.; Remington, S. J. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 2306–2311.

(28) Bublitz, G. U.; King, B. A.; Boxer, S. G. *J. Am. Chem. Soc.* **1998**, *120*, 9370–9371.

(29) Creemers, T. M. H.; Lock, A. J.; Subramaniam, V.; Jovin, T. M.; Völker, S. *Nat. Struct. Biol.* **1999**, *6*, 557–560.

(30) Bell, A. F.; He, X.; Wachter, R. M.; Tonge, P. J. *Biochemistry* **2000**, *39*, 4423–4431.

(31) Voityuk, A. A.; Michel-Beyerle, M. E.; Rösch, N. *Chem. Phys. Lett.* **1997**, *272*, 162–167. *Chem. Phys.* **1998**, *231*, 13−25. *Chem. Phys. Lett*. **1998**, *296*, 269−276.

(32) Weber, W.; Helms, V.; McCammon, J. A.; Langhoff, P. W. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 6177–6182.

(33) Laino, T.; Nifosì, R.; Tozzini, V. *Chem. Phys.* **2004**, *298*, 17–28.

(34) Helms, V.; Winstead, C.; Langhoff, P. W. *J. Mol. Struct.* **2000**, *506*, 179–189.

(35) El Yazal, J.; Prendergast, F. G.; Shaw, D. E.; Pang, Y.-P. *J. Am. Chem. Soc.* **2000**, *122*, 11411–11415.

(36) Marques, M. A. L.; López, X.; Varsano, D.; Castro, A.; Rubio, A. *Phys. Rev. Lett.* **2003**, *90*, 258101−1-4.

(37) Martin, M. E.; Negri, F.; Olivucci, M. *J. Am. Chem. Soc.* **2004**, *126*, 5452–5464.

(38) Sinicropi, A.; Andruniow, T.; Ferré, N.; Basosi, R.; Olivucci, M. *J. Am. Chem. Soc.* **2005**, *127*, 11534–11535.

(39) Vendrell, O.; Gelabert, R.; Moreno, M.; Lluch, J. M. *J. Am. Chem. Soc.* **2006**, *128*, 3564–3574.

(40) Das, A. K.; Hasegawa, J.-Y.; Miyahara, T.; Ehara, M.; Nakatsuij, H. *J. Comput. Chem.* **2003**, *24*, 1421–1431.

(41) Toniolo, A.; Olsen, S.; Manohar, L.; Martínez, T. J. *Faraday Discuss.* **2004**, *127*, 149–163.

(42) Xie, D.; Zeng, X. *J. Comput. Chem.* **2005**, *26*, 1487–1496.

(43) Nemukhin, A. V.; Topol, I. A.; Burt, S. K. *J. Chem. Theory Comput.* **2006**, *2*, 292–299.

(44) Wan, S.; Liu, S.; Zha, G.; Chen, M.; Han, K.; Sun, M. *Biophys. Chem.* **2007**, *129*, 218–223.

(45) Hasegawa, J.-Y.; Fujimoto, K.; Swerts, B.; Miyahara, T.; Nakatsuji, H. *J. Comput. Chem.* **2007**, *28*, 2443–2452.

(46) Bravaya, K. B.; Bochenkova, A. V.; Granovskii, A. A.; Nemukhin, A. V. *Russ. J. Phys. Chem. B* **2008**, *2*, 671–675.

(47) Monkhorst, H. J. *Intern. J. Quantum Chem. Symp.* **1977**, *11*, 421. Dalgaard, E.; Monkhost, H. J. *Phys. Rev. A* **1983**, *28*, 1217–1222.

(48) Foulkes, W. M. C.; Mitas, L.; Needs, R. J.; Rajagopal, G. *Rev. Mod. Phys.* **2001**, *73*, 33–83.

(49) *Molecular Electronic-Structure Theory*; Helgaker, T., Jorgensen, P., Olsen, J., Eds.; John Wiley and Sons, Inc.: Chichester, West Sussex, England, 2000.

(50) Schautz, F.; Filippi, C. *J. Chem. Phys.* **2004**, *120*, 10931–10941.

(51) Schautz, F.; Buda, F.; Filippi, C. *J. Chem. Phys.* **2004**, *121*, 5836–5844.

(52) Cordova, F.; Doriol, L. J.; Ipatov, A.; Casida, M. E.; Filippi, C.; Vela, A. *J. Chem. Phys.* **2007**, *127*, 164111−1-16.

(53) Tapavicza, E.; Tavernelli, I.; Röthlisberger, U.; Filippi, C.; Casida, M. E. *J. Chem. Phys.* **2008**, *129*, 124108−1-18.

(54) Umrigar, C. J.; Toulouse, J.; Filippi, C.; Sorella, S.; Henning, R. G. *Phys. Rev. Lett.* **2007**, *98*, 110201−1-4.

(55) Nightingale, M. P.; Melik-Alaverdian, V. *Phys. Rev. Lett.* **2001**, *87*, 043041−1-4.

(56) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03, Revision C.02*; Gaussian, Inc.: Wallingford CT, 2004.

(57) ADF2007 01, SCM, Theoretical Chemistry, Vrije Universiteit, Amsterdam, The Netherlands. http://www.scm.com (accessed June 15, 2009).

(58) Fonseca Guerra, C.; Snijders, J. G.; te Velde, G.; Baerends, E. J. *Theor. Chem. Acc.* **1998**, *99*, 391–403.

(59) te Velde, G.; Bickelhaupt, F. M.; van Gisbergen, S. J. A.; Fonseca Guerra, C.; Baerends, E. J.; Snijders, J. G.; Ziegler, T. *J. Comput. Chem.* **2001**, *22*, 931–967.

(60) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098–3100. Lee, C. T.; Yang, W. T.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785–789.

(61) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648–5652.

(62) Schipper, P. R. T.; Gritsenko, O. V.; van Gisbergen, S. J. A.; Baerends, E. J. *J. Chem. Phys.* **2000**, *112*, 1344–1352.

(63) Karlström, G.; Lindh, R.; Malmqvist, P.-Å.; Roos, B. O.; Ryde, U.; Veryazov, V.; Widmark, P.-O.; Cossi, M.; Schimmelpfennig, B.; Neogrady, P.; Seijo, L. *Comput. Mater. Sci.* **2003**, *28*, 222–239.

(64) Finley, J.; Malmqvist, P.-A.; Roos, B. O.; Serrano-Andrés, L. *Chem. Phys. Lett.* **1998**, *288*, 299–306.

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

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

(67) MOLPRO, version 2006.1, a package of ab initio programs written by H.- J. Werner, P. J. Knowles, R. Lindh, F. R. Manby, M. Schütz, and others (Cardiff, UK, 2006). http://www.molpro.net (accessed June 15, 2009).

(68) CHAMP is a quantum Monte Carlo program package written by C. J. Umrigar, C. Filippi, and collaborators.

(69) Burkatzki, M.; Filippi, C.; Dolg, M. *J. Chem. Phys.* **2007**, *126*, 234105−1-8.

(70) Filippi, C.; Umrigar, C. J. *J. Chem. Phys.* **1996**, *105*, 213–226. As the Jastrow correlation factor, we use the exponential of the sum of three fifth-order polynomials of the electron-nuclear (e-n), the electron-electron (e-e), and of pure 3-body mixed e-e and e-n distances, respectively. The Jastrow factor is adapted to deal with pseudo-atoms, and the scaling factor κ is set to 0.60 a.u.

(71) 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.

(72) Casula, M. *Phys. Rev. B* **2006**, *74*, 161102(R)-1-4.

(73) Shimomura, O. *FEBS Lett.* **1979**, *104*, 220–222.

(74) Casida, M. E.; Casida, K. C.; Salahub, D. R. *Int. J. Quantum Chem.* **1998**, *70*, 933–941.

(75) Wasserman, A.; Maitra, N. T.; Burke, K. *Phys. Rev. Lett.* **2003**, *91*, 263001-1-4.

(76) Kamlet, M. J.; Abboud, J.-L. M.; Abraham, M. H.; Taft, R. R. *J. Org. Chem.* **1983**, *48*, 2877–2887.

CT900227J