12-2009

Modeling The Fluorescence Of Protein-embedded Tryptophans With Ab Initio Multiconfigurational Quantum Chemistry: The Limiting Cases Of Parvalbumin And Monellin

Sara Pistolesi

Adalgisa Sinicropi

Rebecca Pogni

Riccardo Basosi

Nicolas Ferre

See next page for additional authors

Modeling the Fluorescence of Protein-Embedded Tryptophans with ab Initio Multiconfigurational Quantum Chemistry: The Limiting Cases of Parvalbumin and Monellin

Sara Pistolesi,[†] Adalgisa Sinicropi,[†] Rebecca Pogni,[†] Riccardo Basosi,[†] Nicolas Ferré,[‡] and Massimo Olivucci*[,†,§]

Dipartimento di Chimica, Università di Siena, via Aldo Moro 2, I-53100 Siena, Italy, UMR 6264: Laboratoire Chimie Provence, Universités d'Aix-Marseille I, II et III-CNRS, Faculté de St-Jérôme Case 521, 13397 Marseille cedex 20, France, and Chemistry Department, Bowling Green State University, Bowling Green, Ohio 43403

Received: August 22, 2009

We show that a quantum-mechanics/molecular-mechanics strategy based on ab initio (i.e., first principle) multiconfigurational perturbation theory can reproduce the spectral properties of a tryptophan residue embedded in the contrasting hydrophobic and hydrophilic environments of parvalbumin and monellin, respectively. We show that the observed absorption and emission energies can be reproduced with a less than 3 kcal mol^{-1} error. The analysis of the computed emission energies based on a protein disassembly scheme and protein electrostatic potential mapping allows for a detailed understanding of the factors modulating the tryptophan emission. It is shown that for monellin, where the tryptophan is exposed to the solvent, the fluorescence wavelength is controlled not only by the distribution of the point charges of the protein−solvent environment but also by specific hydrogen bonds and, most important, by the environment-induced change in chromophore structure. In contrast, in parvalbumin, where the chromophore is embedded in the protein core, the structure and emission maxima are the same as those of an isolated 3-methylindole fluorophore. Consistently, we find that in parvalbumin the solvation does not change significantly the computed emission energy.

Introduction

The computer-aided design of unnatural proteins with specific optical properties, such as color and luminescence, represents a complex problem. In these cases, the quantum chemical method employed must be capable to describe both ground and electronically excited states of the protein chromophore. In particular, the computational description of a luminescent (e.g., fluorescent) protein implies the use of methodologies capable of predicting the excited state equilibrium structure and reactivity of fluorophores characterized, even for singlet states, by mixtures of open-shell and charge transfer characters. The ab initio (i.e., first-principle) complete-active-space self-consistent-field (CASSCF) method[1] is a multiconfigurational method offering maximum flexibility for an unbiased (i.e., with no empirically derived parameters and avoiding single-determinant wave functions) description of the electronic and equilibrium structure of the ground and excited states of a molecule. Furthermore, the CASSCF wave function can be readily used for subsequent multiconfigurational second-order perturbation theory[2] computations (CASPT2) of the dynamic correlation energy of each state ultimately allowing for a quantitative evaluation of energy gap between different electronic states.[3,4]

In recent studies[5] we have focused on the prediction of the spectra of proteins featuring cationic or anionic chromophores. In particular, we have shown that an ab initio CASPT2//CASSCF quantum chemical procedure (equilibrium geometries and electronic energies are determined at the CASSCF and CASPT2 levels, respectively) allows, when used as part of a quantum-mechanics/molecular-mechanics (QM/MM) method, for the evaluation of the excitation energy of the retinal protonated Schiff base chromophore[6] (treated quantum mechanically) of rhodopsin (treated using molecular mechanics force fields), of the conjugated bases of the *para*-hydroxy-benzylidene imidazolone fluorophore[7] embedded in the β-barrel scaffold of the green fluorescent protein, and of the *para*-hydroxy-cinnamyl chromophore of the photoactive yellow protein.[8] On the other hand, the simulation of proteins featuring neutral chromophores/fluorophores has, to our knowledge, never been attempted using the same technology.

The fluorescence of tryptophan, due to its 3-methylindole (3-MI_{fluor}) moiety, is sensitive to the molecular environment.[9−13] The 3-MI_{fluor} emission maxima fall in the 308−355 nm[14] range, with the most blue-shifted emissions associated with apolar environments. For this reason, its fluorescence spectrum and intensity have been used to gather structural information and follow protein structural changes such as during folding or denaturation.[15] Due to its spectral properties and wide distribution, tryptophan appears as an ideal candidate for high-level (i.e., "brute force") quantum chemical investigation of a protein-embedded neutral fluorophore. Accordingly, one specific goal of the present study is to gain an atomic-level understanding of the factors controlling the fluorescence wavelength of simple proteins containing a single tryptophan residue with respect to the isolated gas-phase 3-methylindole (3-MI_{gas}). This information can be used to rationalize/predict quantitatively the fluorescence of different proteins or protein mutants where a tryptophan probe has been inserted in certain key positions.[16]

In the following, we use a CASPT2//CASSCF/AMBER/6-31G* QM/MM protocol to compute the tryptophan residue absorption (λ^{a}_{max}) and emission (λ^{f}_{max}) maxima values of two

* Corresponding author. E-mail: olivucci@unisi.it or molivuc@bgsu.edu.
[†] Università di Siena.
[‡] UMR 6264: Laboratoire Chimie Provence.
[§] Bowling Green State University.

10.1021/jp9080993 CCC: $40.75
Published on Web 10/16/2009

very different proteins. These are the F102W mutant of carp parvalbumin[17] (Parv), a calcium-binding protein, and monellin[18] (Mone), a sweet protein extracted from the African serendipity berry *Dioscoreophyllum cumminsii*. While, due to their high cost, CASPT2//CASSCF/AMBER computations cannot presently be used for screening a large set of proteins, the chosen proteins represent different limiting cases. In fact, according to their crystallographic structures, in Parv the single tryptophan residue is located in a substantially hydrophobic cavity that does not contain solvent (water) molecules (see left side of Figure 1). In contrast, in Mone, the tryptophan residue is located on the external surface of the protein in direct contact with the solvent (see right side of Figure 1). The fact that the observed value of λ^{f}_{max} of the two proteins differs by <25 nm (<6.5 kcal mol^{-1} difference in excitation energy) makes them challenging candidates for modeling and analysis.

Below we show that, for both Mone and Parv, the computed λ^{a}_{max} and λ^{f}_{max} values reproduce the observed quantities with a blue-shifted error of <8 nm (equivalent to a ≤2.5 kcal mol^{-1} error). These results compare well with those reported for charged chromophores such as, for instance, the λ^{a}_{max} computed for a set of retinal proteins with a <35 nm (<3.0 kcal mol^{-1}) blue-shifted error. Notice also that the relative λ^{f}_{max} changes were reproduced with a <2.0 kcal mol^{-1} error. Such limited errors allow for an atomic-level analysis of the strongly red-shifted fluorescence of Parv and Mone with respect to 3-MI$_{gas}$. Below we demonstrate that in Parv such λ^{f}_{max} change arises from the stabilization of the ^1L_a emitting state (with respect to the ground state) due, exclusively, to the anisotropy of the *positive* electrostatic potential generated by the protein. On the other hand, the larger λ^{f}_{max} change seen in Mone is due to different effects including a change in the 3-MI$_{fluor}$ geometry (accounting for ca. 50% of the observed red-shift with respect to 3-MI$_{gas}$) and a positive electrostatic potential imposed, cooperatively, by the protein residues and average solvent configuration.

Methods and Models

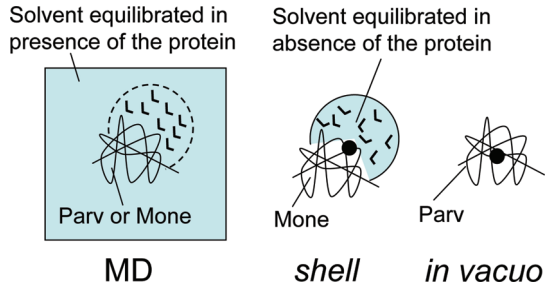
A full description of our QM/MM protocol and protein model building is given in the Supporting Information. Briefly, our QM method is based on a hydrogen link-atom scheme[19] with the frontier placed at the $C_\alpha-C_\beta$ bond of the Trp102 and Trp3 side chains of Parv and Mone, respectively. The selected CASSCF active space comprises the full π-system of 3-MI$_{fluor}$ (ten electrons in nine orbitals). The MM (we use the AMBER force field) and QM segments interact in the following ways: (i) the QM electrons and the full set of MM point charges interact via the one-electron operator, (ii) stretching, bending, and torsional potentials involving at least one MM atom are described by the MM potential, and (iii) QM and MM atom pairs separated by more then two bonds interact via standard van der Waals potentials. CASSCF/6-31G*/AMBER geometry optimization is carried out with the GAUSSIAN03[20] and TINKER[21] programs.

As detailed below, we define two types of protein models featuring different levels of approximation. The first (Parv-MD and Mone-MD) are fully solvated models in which the average solvent configuration is determined via molecular dynamics (MD) equilibration (see left side of Scheme 1). The second is a cruder model that lacks the solvent in Parv (Parv in vacuo) and features a 3-MI$_{fluor}$ shielding shell of solvent molecules equilibrated in the absence of the protein (i.e., featuring the configuration of the pure solvent) in Mone (Mone *shell*) (see right side of Scheme 1). Parv in vacuo and Mone *shell* are analytical tools used to help to determine the factors controlling fluorescence shifts.

The Parv-MD model is based on the F102W mutant and is derived from monomer A of the crystallographic structure deposited in the Protein Data Bank (PDB) archive as file 1B8R,[17] while model Mone-MD is derived from monomer A deposited in the PDB archive as file 1IV7.[18] To get globally neutral models, one sodium and two chloride ions have been added to Parv and Mone, respectively. The average configuration of the solvent surrounding the proteins has been produced according to the following protocol: the PDB structures are embedded in large rectangular boxes ($60 \times 63 \times 53$ and $48 \times 60 \times 53$ Å^3 equivalent to 3569 and 2623 solvent molecules placed within 8 Å from any given Mone and Parv atom, respectively) of TIP3P waters that were energy minimized at the MM level for 2000 steps using the steepest descent method and the AMBER force field. Then, 500 ps MD simulations of the solvent were performed using the Sander module of the Amber 7.0 package[22] with the standard parameters. Both the MM energy minimization and MD simulation were carried out using periodic boundary conditions to simulate the solvent bulk. Coordinates coming from the last frame were used to build the final QM/MM models. This required a CASSCF/6-31G*/AMBER geometry optimization relaxing the coordinates of the 3MI$_{fluor}$ (the full QM subsystem) together with those of the TIP3P water molecules and side chains within 5 Å from any given QM atoms (due to such threshold, no solvent molecules are optimized for Parv). The optimizations have been stopped when the maximum force is <0.003 u.a./bohr and the rms is <0.0005 bohr. The equilibrium structures for the ground state (S_0) and for the first and second singlet excited states (S_1, S_2) were calculated. The S_0 relaxation always produces a structure very close to the crystallographic one. Notice that during the S_1 and S_2 optimization the side chains and solvent molecules surrounding the fluorophore may change position/orientation to adapt to the corresponding excited state charge distribution of 3-MI$_{fluor}$. In Mone-MD, the 39 relaxed solvent molecules interact with the equilibrated fixed outer solvent configuration. The outer solvent extends 16, 11, and 16 Å away from the closest 3-MI$_{fluor}$ atom along the *x*, *y*, and *z* directions of the solvent box (corresponding to more than four layers of water surrounding the fluorophore). In Parv, the solvent box extends 24, 14, and 17 Å away from the fluorophore located at the center of the protein matrix. During the QM/MM calculations, long-range electrostatic effects are not included beyond the solvent box boundaries. Since the 3-MI$_{fluor}$ fluorophores of Parv-MD and Mone-MD are located in different regions of the box we cannot exclude different accuracies in the evaluation of the corresponding excitation energies. Again, notice that the coordinates of the full protein backbone and side chains located >5 Å away from the QM atoms are kept fixed at their crystallographic values (these are considered average values). Due to the excessive computational cost, no second derivative computations could be performed to rigorously determine the nature of the stationary point.

The reduced Parv in vacuo and the Mone *shell* models are derived using the crystallographic structures and QM/MM optimization protocol seen above but feature a different treatment of the solvent environment. Parv in vacuo does not include the solvent at all. In Mone *shell* the solvent is represented by a shell of solvent molecules set up by placing the systems in a rectangular box of TIP3P water molecules positioned within 8 Å from any given atom of the protein using the xleap module of the Amber package.[22] This module provides an initial configuration of the solvent corresponding to a snapshot of MD

SCHEME 1: Type of Models Used for the Construction of the Ground and Excited State QM/MM Equilibrium Structures



equilibrated waters *in the absence of the solute*. To determine a relaxed solvent configuration, this is then minimized for 2000 steps (including counterions) using the steepest descent method (without MD equilibration) while keeping the solute fixed. The resulting configuration provides a solute cavity that, in contrast with the Mone-MD cavity, *is not adapted to the solute point charges*. The final QM/MM model is constructed by discarding all solvent molecules except those forming a solvent sphere of 8 Å radius centered on the $3MI_{fluor}$ chromophore. The equilibrium structures for Mone *shell* is then determined by CASSCF/6-31G*/AMBER geometry optimization to relax the coordinates of the QM chromophore as well as the MM residues and water molecules within 5 Å from any given atom of the $3MI_{fluor}$ moiety. The positions of the remaining solvent molecules (in this model, an ca. 3 Å shield of frozen solvent molecule provides a cage keeping the solvent shell in the correct position and density), protein backbone, and distant residues were kept frozen during QM/MM calculations. We use this model to test the effect of the correct orientation of the solvent molecules surrounding the fluorophore.

In all cases, the residue charges are described by the standard AMBER force field,[23] and thus the residue polarizability or/and dispersion effects are not explicitly treated. As originally pointed out by Warshel,[24,25] a correct QM/MM protein model should include the solvent and account for the solvent and protein polarizability. On the other hand, the effect of the residue polarizability and dispersion on the absolute excitation energy has been shown to be limited (for bR, Warshel et al.[26] estimated an effect of <1500 cm^{-1}; see also the work by Ren et al.,[27] Matsuura et al., and Rajamani et al.[28]) and shall fall into the reported error (however, notice that cancellation effects cannot be excluded). Due to cancellation effects, the error on excitation energy changes will be even smaller.

In the context of the present work, the prepared Mone-MD and Parv-MD structures are assumed to provide acceptable representation of the average environment of the chromophore/fluorophore (i.e., given the high >40% content of solvent in the crystals—see the PDB files—we assume that the crystallographic structures provide a suitable protein average structure in solution, and possible large-amplitude low-frequency fluctuations affecting the average structure are assumed to have a small effect on the observed λ^a_{max} and λ^f_{max} values). Furthermore, within our fixed charge model, the excitation energy should be mainly determined from the residues belonging to the chromophore cavity. This assumption has been assessed[29] by evaluating the λ^a_{max} of rhodopsin with a full protein model and with a reduced protein model consisting of the 27 residues surrounding the retinal chromophore and two crystallographic waters. It is shown that the two λ^a_{max} values differ by <1 kcal mol^{-1}.

For all protein models, at the S_0 equilibrium geometries a CASPT2 computation is carried out, using the MOLCAS-6 program,[30,31] to evaluate the vertical excitation energy for the ($S_0 \rightarrow S_1$ and $S_0 \rightarrow S_2$) transitions (assumed to match the energy of the corresponding λ^a_{max} values) and the associated oscillator strength. The emission λ^f_{max} values from the first and second excited states ($S_1 \rightarrow S_0$ and $S_2 \rightarrow S_0$ transitions) are evaluated by computing the S_1-S_0 and S_2-S_0 vertical energy gaps via CASPT2 computations using a 0th-order three-root state average CASSCF wave function at the S_1 and S_2 equilibrium geometry, respectively.

Again, notice that while the AMBER charges account for S_0 polarization effects in a mean-field way[32] no polarizable residue is included in the protein model. The same charges are used for the excited state computations without introducing an ad hoc dielectric constant. Also, notice that, in this work, there is no scaling of the computed CASPT2//CASSCF/AMBER excitation energies.

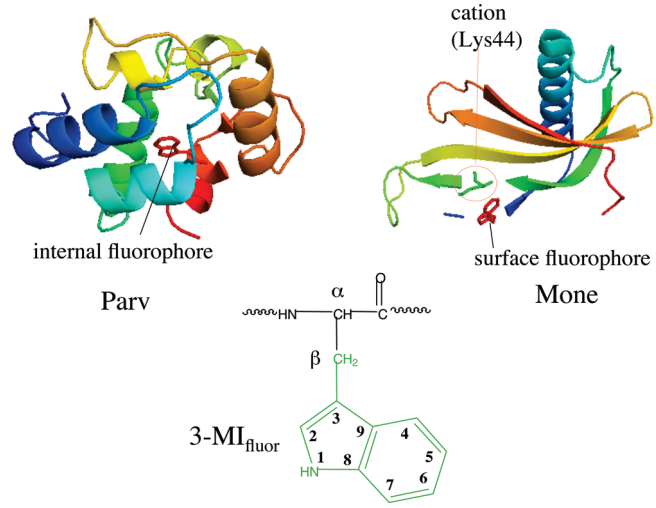


Figure 1. Top. View of parvalbumin (left) and monellin (right) structures. The positions of the 3-methylindole (red) and of the charged residue of monellin (Lys44) are highlighted. Bottom. Structure of the 3-methylindole (green) fluorophore.

Results and Discussion

The changes in the tryptophan fluorescence as a function of the protein structure are a manifestation of the diversity of protein environments.[15] Since the stiff bicyclic framework of the $3MI_{fluor}$ structure does not allow for significant deconjugation of its π-system, the observed spectral change is expected to originate mainly from a change in the electrostatic potential acting on the fluorophore centers. In other words, as pointed out by Callis and co-workers,[14] the λ^f_{max} value and intensity must depend on the "electrostatic potential landscape". Most important, also the fluorescence lifetime and its decay dynamics are affected by such a landscape.[13] In this context, the possibility to employ ab initio (i.e., unbiased) multiconfigurational quantum chemistry opens up new perspectives. Indeed, in the past, CASPT2//CASSCF photochemical reaction path computations[33] have shown that (nonadiabatic) hydrogen and charge (electron) transfer mediated by conical intersections provide the main channel of lifetime modulation (e.g., through fluorescence quenching) of synthetic fluorophores such as azoalkanes and ketones.[34−36] Thus, the development of CASPT2//CASSCF/AMBER protocols for the evaluation of excited state reaction paths[6] and trajectories[37] should allow for the extension of such studies to protein fluorophores.

As a preliminary step toward the investigation of the tryptophan fluorescence and its decay dynamics in proteins with

TABLE 1: Computed and Observed Absorption Maxima (λ_{max}) and Computed Change in Dipole Moments ($\Delta\mu$) for the Vertical Transitions of 3-MI$_{gas}$, Parv, and Mone[a]

protein model (equilibrium struct.)	calcd λ_{max} (nm)	obsd λ_{max} (nm)	$\Delta\mu$ (Debye)
GS-3-MI$_{gas}$ ($S_0 \rightarrow S_1$)	277 (103.2) [0.02]	282 (101.3) {6.9} ^1L_b = [0.045]	0.18
GS-3-MI$_{gas}$ ($S_0 \rightarrow S_2$)	249 (115.0) [0.09]	^1L_a = [0.123]	2.08
^1L_a-3-MI$_{gas}$ ($S_2 \rightarrow S_0$)	295 (97.1) [0.11] *96.98*	295 (96.9)	5.85
GS-Parv-MD ($S_0 \rightarrow S_1$)	279 (102.7) [0.02]	275 (103.9) {5.5}	0.46
GS-Parv-MD ($S_0 \rightarrow S_2$)	257 (111.5) [0.08]		
^1L_a-Parv-MD ($S_1 \rightarrow S_0$)	308 (92.9) [0.10] *92.14*	316 (90.4)	5.49
GS-Mone-MD ($S_0 \rightarrow S_1$)	279 (102.7) [0.02]	277 (103.2) {14.7}	0.56
GS-Mone-MD ($S_0 \rightarrow S_2$)	261 (109.7) [0.07]		5.64
^1L_a-Mone-MD ($S_1 \rightarrow S_0$)	336 (85.2) [0.09] *84.31*	339 (84.3)	5.61

[a] Excitation energies (kcal mol^{-1}) are given in parentheses. Oscillator strengths are given in square brackets. The change in oscillator strengths can be compared with the variation in the observed extinction coefficients (10^{-3} M^{-1} cm^{-1}) given in curly brackets. The values of CASPT2//CASSCF/ANO-S (C,N[4s3p1d]/H[2s]) energies, computed at the corresponding CASSCF/6-31G* optimized geometries, are given in italics. Computed spectral parameter data for the parent gas-phase indole (not reported below) using the same protocol and the (C,N[3s2p1d]/H[2s]) ANO-S basis are reported in ref 38.

multiconfigurational quantum chemistry tools, here we focus on the origin of the λ^f_{max} value. For other photoactive proteins, four different molecular factors have been proposed to play a role at the molecular level: (i) the increased or decreased conjugation of the π-system;[34,35] (ii) the placement of charged, polarized, or polarizable groups close to the fluorophore unit;[36] (iii) the contact with the solvent molecules;[39] and (iv) the chromophore orientation with respect to the protein cavity.[5] Factor ii and iii are usually considered dominating.

Past work employing an INDO/S-CIS semiempirical quantum chemical protocol coupled with the force field CHARMM on several proteins containing a single tryptophan residue has shown that it is possible, using a specific scaling factor for the fluorophore charges, to reproduce systematically the fluorescence wavelength within a <15 nm error (with respect to the available experimental data). On the basis of these data, it was proposed that, in Mone, the λ^f_{max} value is controlled by the polarization of the indole ring by the electric fields of both the solvent (water) and protein charges, while in Parv, the λ^f_{max} is mainly controlled by the geometrical polarization of the external water molecules.[14] However, a quantitative analysis of factors i−iv based on the results of ab initio multiconfigurational quantum chemistry has never been reported.

In Table 1, we report the excitation energies, evaluated at the ground state (S_0) equilibrium structure, and the emission energies, evaluated at the ^1L_a equilibrium structure, for the Parv-MD and Mone-MD models together with those of the reference system 3-MI$_{gas}$. In their S_0 equilibrium geometry, the proteins display a slightly lower ^1L_b state (see the Supporting Information). In contrast, in the ^1L_a excited state geometry, S_1 always corresponds to the ^1L_a state, thus displaying charge transfer character (i.e., a larger oscillator strength and dipole moment change). Such $^1L_a/^1L_b$ inversion does not occur in 3-MI$_{gas}$ where the covalent ^1L_b state is always lower (12 kcal mol^{-1} at GS-3-MI$_{gas}$). The computed emission energies reproduce the observed values with systematically blue-shifted errors of ≤2.5 kcal mol^{-1} (see Figure 2A). The previous (semiempirical) INDO/S-CIS based study by Callis predicted, for Mone and Parv, red-shifted λ^f_{max} 6.7 and 1.3 kcal mol^{-1} of the observed values, respectively. These errors were reduced to 3.0 and 0.7 kcal mol^{-1} after scaling of the fluorophore charges. Our unscaled CASPT2//CASSCF/AMBER QM/MM protocol yields, for the same quantities, 0.9 and 2.4 kcal mol^{-1} blue-shifted errors, respectively.

As described in the Methods and Models section, Parv in vacuo and Mone *shell* are cruder Parv and Mone models. Parv in vacuo has no solvent. The errors in the predicted λ^a_{max} and

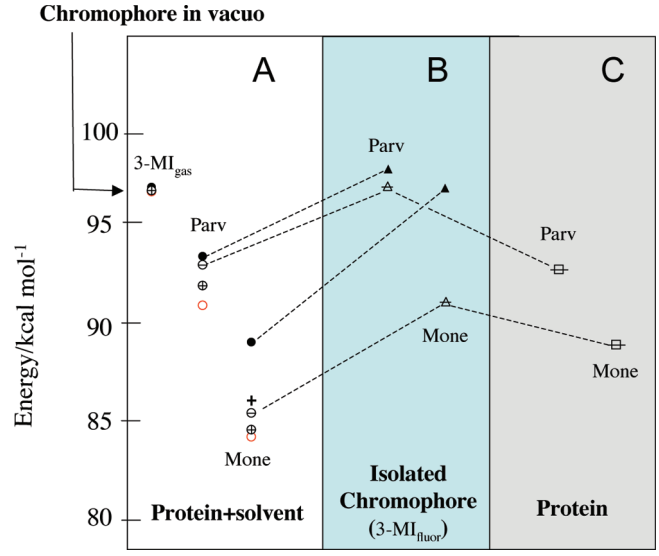


Figure 2. Analysis of the CASPT2//CASSCF/6-31G*/Amber emission energies (equilibrium ^1L_a structures) of 3-MI$_{gas}$, Parv, and Mone models. (A) Full circles: Parv in vacuo, Mone *shell*, 3-MI$_{gas}$ (see Methods and Models). Crossed circles: ANO-S emission energies of 3-MI$_{gas}$, Parv-MD, and Mone-MD. Half open circles: Fully solvated models Mone-MD and Parv-MD. Cross: Mone-MD deprived of the solvent bulk (Mone-MD-shell). Open circles: experimental values. (B) Isolated chromophores taken with their protein equilibrium geometries. Full triangles: Parv in vacuo and Mone *shell*. Half open triangles: Parv-MD and Mone-MD. (C) Protein embedded chromophores (i.e., with removed solvent) taken with the Mone-MD and Parv-MD equilibrium geometries. Half open squares: Mone-MD and Parv-MD.

λ^f_{max} values computed for such a model are similar to the one found for Parv-MD. This reflects a limited influence of the solvent on the electronic structure of the protein-embedded fluorophore. In contrast, the Mone *shell* model displays a significantly larger blue-shifted error (4.5 kcal mol^{-1}) with respect to the values predicted for Mone-MD. Consistently with the Parv in vacuo model, this effect does not come from the model lack of solvent bulk. In fact, removal of the solvent bulk in Mone-MD while keeping a solvent shell of waters within 5 Å from any 3-MI$_{fluor}$ atom generates a model (here indicated as Mone-MD-shell) that yields an excitation energy that is only 1 kcal mol^{-1} larger than the fully solvated system (see the Mone-MD-shell value in Figure 2A). In other words, a limited solvent shell with the correct configuration will reproduce most of the red-shifting solvent effect. The larger λ^f_{max} error seen in the Mone *shell* is due to a solvent average configuration that is not

adapted to the solvent-exposed 3-MI_{fluor} and surrounding Mone charge distribution. In conclusion, the effect of solvent is, as expected, limited in Parv where the fluorophore is embedded in a hydrophobic cavity but large (e.g., showing a 3-fold error increase with respect to the observed λ^{f}_{max} value) in Mone where the fluorophore is placed at the surface of the protein and in direct contact with the solvent. In the next section, we will provide a rationalization for these effects on the basis of the electrostatic potential acting on the 3-MI_{fluor} centers.

The error coming from the use of the 6-31G* basis with a correlated CASPT2//CASSCF wave function has been investigated by comparing the excitation energies computed for 3-MI_{gas} with ANO-S data from the literature (see Table 1). We also recomputed the λ^{f}_{max} of the full models Mone-MD and Parv-MD using the same ANO-S (C,N[4s3p1d]/H[2s]) basis set with respect to the 6-31G* basis set. This yields a smaller blue-shifted error (see Figure 2A and the Supporting Information) indicating that, as expected, a better basis increases the computational accuracy. However, since we focus on λ^{f}_{max} *changes* and to avoid excessive computational costs, below we focus on the 6-31G* results.

Effect of the Fluorophore Environment. The limited, *relative* error in excitation energies computed for Mone-MD, Parv-MD, and 3-MI_{gas} prompts for an analysis of the factors determining the λ^{f}_{max} values. Accordingly, the emission energy of each model is compared with that of the isolated fluorophore taken with its protein-optimized geometry (Figure 2B) and with that of protein deprived of the solvent (Figure 2C).

Inspection of the data in Figure 2B shows that in Parv the change in 3-MI_{fluor} geometry cannot be responsible for the observed red-shifted emission. In fact, the corresponding 3-MI_{fluor} has an excitation energy that is only slightly higher that the reference quantity (i.e., 3-MI_{gas}). This is obviously due to a limited protein-induced change in fluorophore excited state geometry with respect to the gas phase. Such a conclusion is confirmed by the data in Figure 3A pointing to a <0.015 Å difference between the geometrical parameters of 3-MI_{gas} and of the Parv-MD fluorophore in both the ground and excited states. Furthermore, the same quantity is close to the excitation energy computed for the fluorophore of Parv in vacuo (see Figure 2B) demonstrating that the solvent has a small effect on the 3-MI_{fluor} structure (see Figure 3A). This conclusion is further supported by the data in Figure 2C where we show that removal of the solvent from Parv-MD does not significantly change the excitation energy (i.e., this value is very close to the excitation energy of Parv in vacuo seen in Figure 2A).

In contrast to Parv, the data of Figure 2B show that, in Mone, the isolated fluorophore must undergo a large structural change. In fact, the corresponding excitation energy is significantly lower with respect to 3-MI_{gas}. Again, this is consistent with the structure of Figure 3A that shows large deviations of the ground state and excited state 3-MI_{fluor} structures from the 3-MI_{gas} reference. Comparison of the Mone-MD excitation energies in Figure 2A and 2B reveals that a decrease >50% of the excitation energy with respect to 3-MI_{gas} is due to a change in the fluorophore structure. This change is very sensitive to the structure of the solvent shell surrounding 3-MI_{fluor} (or, in other words, to the solvent average configuration model). Indeed, the excitation energies of the Mone-MD and Mone *shell* fluorophores in Figure 2B are very different reflecting the limited change in the structure of the Mone *shell* fluorophore with respect to 3-MI_{gas} (see Figure 3A). A correct solvent shell configuration also induces the structural change that leads to an enhanced dipole moment change and charge separation (see the Supporting Information). Comparison of the Mone data in Figure 2A and 2C indicates that the point charges both of the protein and of the solvent contribute, concurrently, to further decrease the excitation energy of the Mone fluorophore. In conclusion, the contributions of factor i and iii are not critical for Parv but are both important when the fluorophore is in contact with the solvent (as in Mone). The comparison between the Mone-MD and Mone *shell* data in Figure 2B indicates that the structure change of 3-MI_{fluor} is due to factor iii.

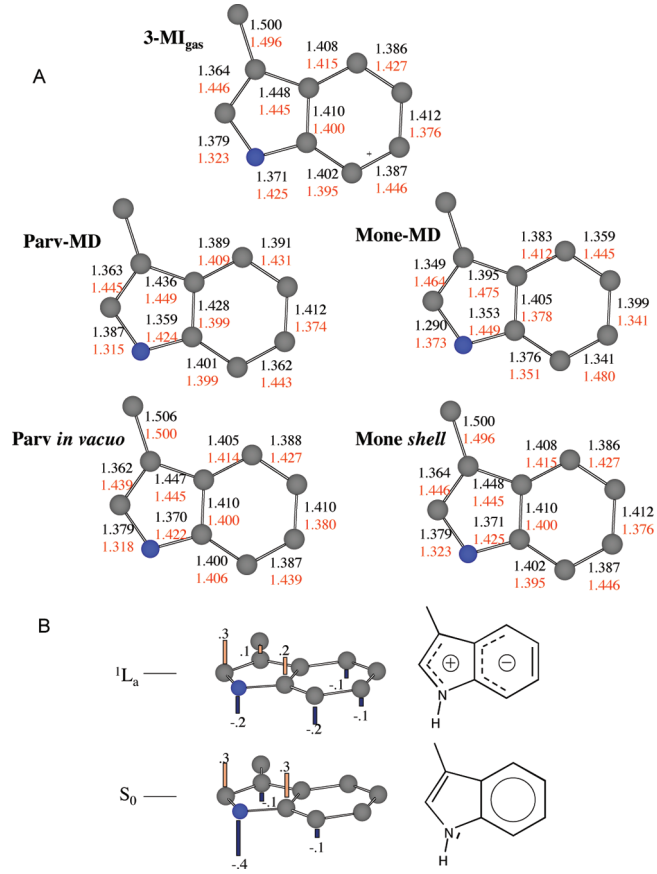


Figure 3. (A) Ground (black) and ^1L_a (red) equilibrium structure parameters for 3-methylindole and the 3-methylindole fluorophore of the Mone and Parv protein models. (B) Charge distribution of the ground (S_0) and excited (^1L_a) state of 3-MI_{gas}. Representative resonance formulas are also displayed.

A common basis for the discussion of the effects of the protein residues and solvent molecules is provided by the charge transfer nature of the spectroscopic state of tryptophan (see Figure 3B for the case of 3-MI_{gas}). In fact, in 3-MI_{gas} 38% of negative charge (π-electron density), originally located on the pyrrole moiety, is shifted toward the benzene ring upon the $S_0 \rightarrow {}^1L_a$ transition. Thus, an electrostatic potential stabilizing the positive charge on the ^1L_a pyrrole moiety or stabilizing the negative charge on the ^1L_a benzene moiety will result in a decreased $^1L_a \rightarrow S_0$ emission energy.[40] Notice that while the ^1L_a equilibrium structure of the chromophore remains planar an extensive bond length rearrangement occurs to accommodate the charge transfer (see Figure 3A). As already mentioned above, both the entity of the charge transfer and the bond readjustment pattern are different in the systems investigated here. The Parv fluorophore shows changes very close to those of 3-MI_{gas}, while as discussed above, in Mone these changes are different.

According to Figure 4A, the anisotropic electrostatic potential (generated from the residue and solvent point charges) acting on the chromophore centers of Parv-MD and Mone-MD is

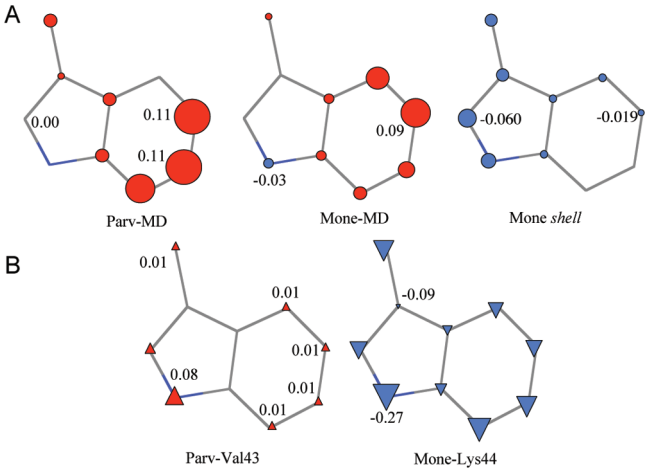


Figure 4. (A) Electrostatic potential on the carbon and nitrogen 3-MI_{fluor} centers of Parv (Parv-MD model) and Mone (Mone-MD and Mone *shell* models). Full circles indicate a negative (blue) or positive (red) potential. (B) 3-MI_{fluor} change in electrostatic potential for Parv-MD and Mone-MD after removal of the Val43 and Lys44 residue charges, respectively. Full triangles indicate changes toward negative (tip down) and positive (tip up) values. On each structure, the numerical values represent the highest and lowest values of the corresponding quantity.

positive and localized on the benzene ring. This potential must thus decrease the S_1-S_0 energy gap relative to the isolated chromophore consistently with the computed λ^f_{max} red-shift. In Parv, the electrostatic potential is due exclusively to the protein charges.

Comparison of the Mone-MD and Mone *shell* electrostatic potential in Figure 4A shows that this is very sensitive to the configuration of the solvent shell surrounding 3-MI_{fluor}. If this configuration is not adapted to the protein and fluorophore charges (i.e., does not correctly represent the average configuration of the solvent), the potential acting on the QM atoms is negative and not positive. Consistently, with the excitation energy and geometrical structure analysis given above, this potential still leads to a (limited) red-shift since it is mainly localized on the pyrrole unit where a positive charge developed on the excited state (see Figure 3B).

Effect of Specific Residues and Solvent Molecules. To assess the effect of specific side chains and water molecules, we focus on the Parv-MD and Mone-MD models. As shown in Figure 5A, Mone features a positively charged residue (Lys44) in the close vicinity of the fluorophore. The effect of the Lys44 on the emission energy can be understood on the basis of the charge transfer nature of the ^1L_a state (see Figure 3B) and position of the Lys44 chain. Since ^1L_a features a more positive pyrrole moiety and a more negative benzene ring with respect to S_0, the Lys44 charge, that is located on top of the benzene moiety, must stabilize the excited state relative to the ground state.

The electrostatic potential acting on the centers of the Mone fluorophores in the absence of the Lys44 charges (Mone-Lys44) is clearly less positive on the phenyl moiety and becomes very negative on the pyrrole moiety (see Figure 4B). This unveils a counterbalancing effect of Lys44 residue charges and solvent shell. To explain this finding (i.e., a negative potential induced by the removal of the positive Lys44 charge) one must focus on the fact that in Mone the 3-MI_{fluor} moiety and solvent shell are adapted to the protein environments. Since the water molecules have a large dipole moment their average orientation will be determined by the interaction with the protein partial or

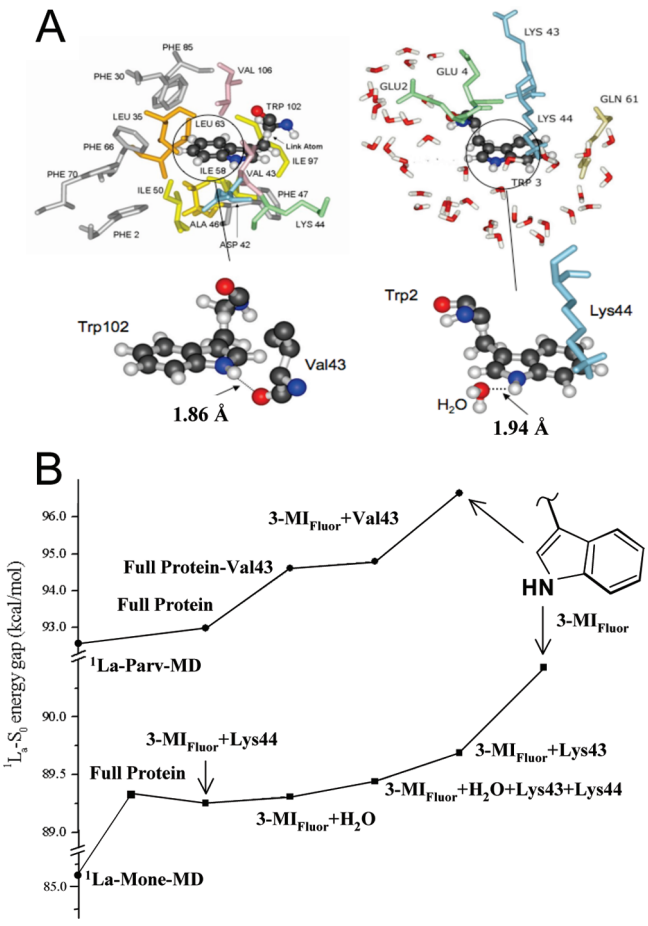


Figure 5. (A) Details of the region surrounding the fluorophore of the computational models Parv-MD (left) and Mone-MD (right). The configuration of the first solvent layer surrounding the Mone fluorophore is displayed (see the Supporting Information for details). The closest solvent molecule forms a hydrogen bond with the fluorophore N−H group. For Parv, the relationship between the fluorophore (Trp102) and one valine residue (Val43) is also given. The fluorophore is not in contact with the solvent, and the closest solvent molecule is 6.28 Å from the C2−H group of the fluorophore. (B) Analysis of the excitation energies of Parv-MD (full circles) and Mone-MD (full squares) at the corresponding ^1L_a equilibrium structures. Each fragment is taken with the geometry optimized for the full protein.

fully charged groups.[14] Thus, in Mone, the positively charged Lys44 residue, which, with respect to 3-MI_{fluor}, is located opposite to the solvent (see the right structure in Figure 5A), will presumably orient the water molecules of the first solvation shell (also through enhanced polarization of the fluorophore π-system) in such a way to point their negatively charged oxygens toward 3-MI_{fluor}. This specific orientation leads to a negative potential (with the largest intensity on the positively charged pyrrole moiety). The $^1L_a-S_0$ gap of the Mone-MD-Shell model with respect to Mone-MD indicates that only the solvent molecules in the close vicinity of the fluorophore are responsible for the described solvent effect and not the bulk.

In Figure 5B we report the results of an analysis of the effect of selected residues and solvent molecules on the emission maximum of the 3-MI_{fluor} of Mone. It is apparent that the Lys44 residue has a red-shifting effect similar to that induced by the full protein. It is also apparent that both the Lys44 residue alone and a specific hydrogen bonded water molecule (linked to the N−H bond of the pyrrole moiety) alone have similar red-shifting effects on 3-MI_{fluor}. It is interesting to see that when both effects are present one still gets a similar red-shift pointing to a different

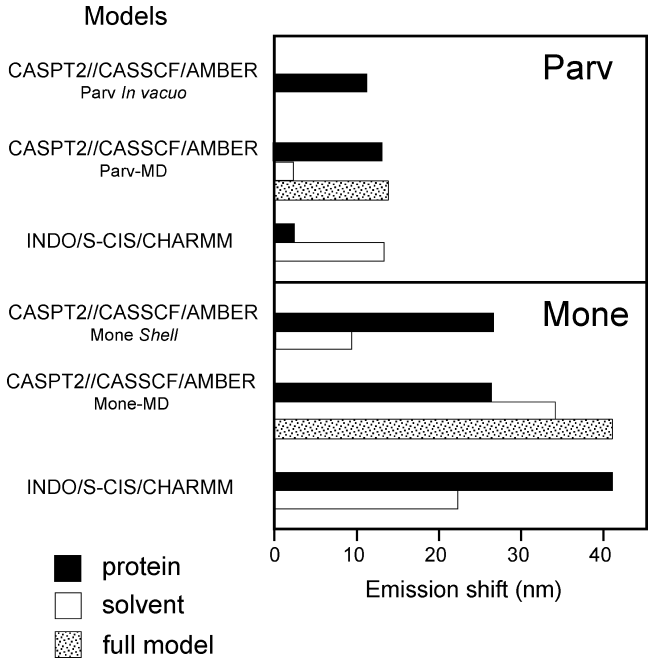


Figure 6. Comparison between predicted emission wavelength red-shift (with respect to in vacuo 3-MI) for Parv and Mone using ab initio CASPT2//CASSCF/AMBER (this paper) and the semiempirical INDO/S-CIS/CHARMM QM/MM protocols (see ref 14). The white and black horizontal bars represent the wavelength change induced by the solvent or protein charges, respectively (e.g., the black bars in Parv-MD and Mone-MD correspond to the data of Figure 2C).

and opposite sign of the electrostatic potential projected on the chromophore. Because of the different sign of the potential, the stabilizing effects partially cancel each other. This seems in line with the picture given above on the basis of the potentials of Figure 4. On the other hand, it is clear that other residues such as, for instance, Lys43 must contribute to the total red-shift with respect to bare 3-MI_{fluor}. As shown in Figure 5B, this effect is smaller due to the largest distance of this residue from the phenyl moiety.

In conclusion, from a mechanistic point of view, our analysis unveils that the dipole of one N−H hydrogen-bonded water molecule and the positive charge of the Lys44 residue are mainly responsible for the low emission energy of tryptophan in Mone. The water dipole stabilizes the positive charge on the pyrrole moiety, while the Lys44 residue stabilizes the negative charge on the benzene moiety of the ^1L_a state of 3-MI_{Fluor}. This analysis reveals the complex origin of the total electrostatic potential that cannot be disentangled in simple additional contributions.

A similar analysis can be carried out for Parv. In this case, we detected a large effect (a ≈2 kcal mol^{-1} red-shift) associated with the Val43 residue of the fluorophore cavity. The remaining cavity residues have, taken individually, a much smaller effect. However, as shown in Figure 5B (see the "Full Protein-Val43" value), these induce, cooperatively, a ≈3 kcal mol^{-1} red-shift. The origin of the Val43 effect has been determined by inspection of the ^1L_a structure of Parv reported in Figure 5A (left). From such a structure, it is apparent that the backbone carbonyl group of Val43 points toward the N−H group of 3-MI_{Fluor} and forms a hydrogen bond. Thus, the computed decrease in emission energy originates from a stabilization of the positive charge on the pyrrole ring that, in turn, induces a stabilization of the excited state with respect to the ground state. Once again, the decomposition of the solvated Parv suggests that the external solvent does not play a key role in controlling λ^f_{max}. Such control is due to the residues in the fluorophore cavity.

The idea that the residues (or solvent molecules) in direct contact or close to the chromophore/fluorophore control the optical properties has also been investigated for the visual photoreceptor rhodopsin.[5] CASPT2//CASSCF/AMBER excitation energy computations demonstrate that only the residues in the chromophore cavity affect significantly the excitation energy and therefore the absorption wavelength.[29] In our Parv model, the electrostatic potential acting on the fluorophore is again positive (see Figure 4B). Therefore, the effect of the hydrogen-bonded Val43 residue must be such to increase the positive potential at the level of the benzene moiety or decrease the positive potential at the level of the indole moiety. As shown in Figure 4B, removal of Val43 leads to a more negative potential on the N−H region of indole. Thus, it is the latter effect to occur in Parv.

Conclusions

The work of Callis and co-workers[14] has established that a QM/MM protocol based on semiempirical quantum chemical methods and a suitable fluorophore charge scaling can successfully reproduce the fluorescence maxima of 19 tryptophan-containing proteins with a ≈15 nm error. In that study, all protein−solvent systems (with a fixed ^1L_a reference geometry for the fluorophore moiety) were modeled via a 30 ps trajectory and the emission maxima determined by averaging the emission energies of many snapshots. While such a protocol cannot be used with state-of-the-art ab initio QM methods, in the past we have provided evidence that both absorption and emission can be simulated using the CASPT2//CASSCF/AMBER protocol. In particular, using a crystallographic structure and a suitably prepared solvent box to model the average protein−solvent environment,[6,7] it has been shown that the absorption or fluorescence of anionic and cationic biological chromophores/fluorophores can be reproduced within <5 kcal mol^{-1}.

Above, the CASPT2//CASSCF/AMBER protocol has been used to study the spectroscopy of two very different protein-embedded tryptophans as examples of neutral biological fluorophores. As shown in Figure 2, and in spite of the absence of empirical parameters, the ab initio CASPT2//CASSCF level allows us to reproduce the observed changes in λ^f_{max} value (i.e., with respect to 3-MI) with errors smaller than those obtained using not scaled semiempirical methods. Most important, the error is of the same magnitude as that found in proteins containing charged chromophores. Therefore, our study extends the validity of the protocol and paves the way to first-principle simulations of the photophysics of protein-embedded fluorescent probes.

We believe that the results presented above will have an impact on future studies of the fluorescence lifetime and decay dynamics of tryptophan. Indeed, previous mechanistic studies of the fluorescence decay dynamics of synthetic fluorophores demonstrate that the decay may occur via aborted chemical or electron transfer reactions mediated by conical intersection channels. This fact calls for the use of multiconfigurational quantum chemical theories (e.g., CASSCF) where real (non-avoided) crossings between potential energy surfaces of the same spin multiplicity are properly represented.[34−36]

Consistently with previous work, we have confirmed that in Parv the environmental effects on the λ^f_{max} value are mainly due to stabilization of the emitting charge-transfer state. On the other hand, it has been possible to unveil that, in contrast to semiempirical studies,[14] the ab initio CASPT2//CASSCF treatment points to a minor effect of the external solvent on the λ^f_{max} change when a fluorophore is embedded in a substantially

hydrophobic protein matrix. In contrast, we have shown that when the fluorophore is in contact with the solvent (e.g., in Mone) the solvent contribution to the λ^{f}_{max} change is very important and also affects the fluorophore structure.

The red-shift values induced by the protein and solvent charges are summarized in Figure 6. For Parv, the computed CASPT2//CASSCF/AMBER and "scaled-charge" INDO/S-CIS/CHARMM emission shifts have very different magnitudes. In fact, the semiempirical data point to a large and dominating effect of the solvent surrounding the protein (i.e., not in contact with 3-MI_{fluor}). In contrast, the ab initio data point to a dominating effect of the protein (consistently with the red-shift value of the Parv in vacuo unsolvated model). On the other hand, for Mone, the ratio between the solvent and protein effects is more balanced. However, in this case the INDO/S-CIS/CHARMM emission points to a protein dominating effect, while the CASPT2//CASSCF/AMBER protocol yields a solvent effect that is larger that the protein effect. Of course, as shown in the figure (see full model bars), these effects are far from being additive as the electrostatic field imposed by the solvent is modified by that of the protein, and this depends on the protein and solvent spatial locations with respect to the fluorophore.

At least three of the specific i−iv factors mentioned above have been found to affect the fluorescence color in Mone and Parv. The first mechanism corresponds to factor i. In fact, in Mone, the bond-length pattern of 3-MI_{fluor} is very different from that of the reference gas-phase fluorophore. It seems that the change is induced by the solvent shell structure. A second important factor is factor ii. This is related to the effect of specific and strategically located protein cavity residues. In Mone, the Lys44 residue with its positive charge placed above the benzene moiety and in Parv the Val43 residue that forms hydrogen bonds with the pyrrole moiety of the fluorophore constitute clear examples. The impact of specific interactions is estimated of the order of 2 kcal mol^{-1} (see Figure 5B). Of course, cooperative effects in the same category (i.e., contributions coming from many fractional charges) may have a considerable weight. The third mechanism is related to the direct electrostatic interaction with the solvent and corresponds to factor iii. It is found that this factor is coupled with factor i and ii. In fact, in Mone both the fluorophore π-system polarization and the geometrical reorganization of the first solvent shell seem to account for part of the red-shifting effect. On the other hand, interactions with specific solvent molecules (e.g., a single water molecule in Mone) account for the emission energy decrease.

In the near future, the development of accurate tools for the simulation of the changes in tryptophan fluorescence and fluorescence lifetimes as a function of the residue position in a peptide backbone may constitute one important tool for the design of unnatural proteins with wanted properties. These tools will have to be potent enough to be able to describe the mechanism of internal residue or solvent mediated quenching. In particular, they should be able to deal with large and rapid changes in the structure and electronic wave function that often characterize photoinduced aborted photochemical or electron transfer reactions. Presently, it is accepted that these processes imply the evolution toward a conical intersection funnel connecting the emitting excited state to the ground state or to a lower nonemitting excited state (e.g., conical intersections are, by definition, regions of rapid change in electronic structure). The investigation of these processes calls for unbiased quantum chemical methods featuring highly flexible wave functions and for the tools of the emerging field of computational photochemistry.[33] Above we have provided evidence that, for two very different proteins, the use of QM/MM strategy based on multiconfigurational second-order perturbation theory can potentially give access to a quantitative mapping of the excited state reaction paths and, in turn, resolve the atomic-level mechanism controlling the fluorescence quenching/decay.[9,11−13,41] Recent work has established that the evaluation of realistic time scales (including the biexponential character of such processes[42]) via scaled-CASSCF/AMBER trajectories is becoming a reality.[37] Given the steady increase in computer performances, the continuation of such exploratory research effort, in our and other laboratories, appears to open new perspectives for the future of de novo light-sensitive protein design.

Acknowledgment. Funds have been provided by the Università di Siena (Progetto di Ateneo 02/04) and the Fondazione Monte dei Paschi di Siena. We thank CINECA for granted calculation time. M.O. is grateful to the Center for Photochemical Sciences and the School of Arts & Sciences of the Bowling Green State University for start-up funds.

Supporting Information Available: The QM/MM scheme, models, solutions models, tables, and optimized structures. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Roos, B. O. Advances in Chemical Physics. In *Ab Initio Methods in Quantum Chemistry - II*; Lawley, K. P., Ed.; Wiley & Sons: Chincester, 1987; pp 399−445.
(2) Andersson, K.; Malmqvist, P.-Å.; Roos, B. O. *J. Chem. Phys.* **1992**, *96*, 1218–1226.
(3) Garavelli, M.; Celani, P.; Bernardi, F.; Robb, M. A.; Olivucci, M. *J. Am. Chem. Soc.* **1997**, *119*, 6891–6901.
(4) Roos, B. O.; Fülscher, M. P.; Malmqvist, P.-A.; Merchan, M.; Serrano-Andrés, L. *Quantum Mechanical Electronic Structure Calculations with Chemical Accuracy*; Langhoff, S. R., Ed.; Kluwer Academic Publishers: Dordrecht, The Netherlands, 1995; pp 357−438.
(5) Coto, P. B.; Strambi, A.; Ferré, N.; Olivucci, M. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 17154–17159.
(6) Andruniów, T.; Ferré, N.; Olivucci, M. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 17908–17913.
(7) Sinicropi, A.; Andruniow, T.; Ferré, N.; Basosi, R.; Olivucci, M. *J. Am. Chem. Soc.* **2005**, *127*, 11534–11535.
(8) Coto, P. B.; Martí, S.; Oliva, M.; Olivucci, M.; Merchán, M.; Andrés, J. *J. Phys. Chem. B* **2008**, *112*, 7153–7156.
(9) Chen, Y.; Barkley, M. D. *Biochemistry* **1998**, *37*, 9976–9982.
(10) Engelborghs, Y. *J. Fluoresc.* **2003**, *13*, 9–16.
(11) Zhang, L.; Wang, L.; Kao, Y.-T.; Qiu, W.; Yang, Y.; Okobiah, O.; Zhong, D. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 18461–18466.
(12) Qiu, W.; Li, T.; Zhang, L.; Yang, Y.; Kao, Y.-T.; Wang, L.; Zhong, D. *Chem. Phys.* **2008**, *350*, 154–164.
(13) Qiu, W.; Kao, Y.-T.; Zhang, L.; Yang, Y.; Wang, L.; Stites, W. E.; Zhong, D.; Zewail, A. H. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 13979–13984.
(14) Vivian, J. T.; Callis, P. R. *Biophys. J.* **2001**, *80*, 2093–2109.
(15) Eftink, M. R. *Methods of Biochemical Analysis*; Suelter, C. H., Ed.; John Wiley & Sons: New York, 1991; Vol. 35.
(16) Mascotti, D. P.; Lohman, T. M. *Biochemistry* **1997**, *36*, 7272–7279.
(17) Moncrieffe, M. C.; Juranic, N.; Kemple, M. D.; Potter, J. D.; Macura, S.; Prendergast, F. G. *J. Mol. Biol.* **2000**, *297*, 147–163.
(18) Somoza, J. R.; Jiang, F.; Tong, L.; Kang, C. H.; Cho, J. M.; Kim, S. H. *J. Mol. Biol.* **1993**, *234*, 390–404.
(19) Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1986**, *7*, 718–730.
(20) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A.; 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.; 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 B.04; Gaussian, Inc.: Pittsburgh PA, 2003.

(21) Ponder, J. W.; Richards, F. M. *J. Comput. Chem.* **1987**, *8*, 1016–1024.

(22) Case, D. A.; Pearlman, D. A.; Caldwell, J. W.; III, T. E. C.; Wang, J.; Ross, W. S.; Simmerling, C. L.; Darden, T. A.; Merz, K. M.; Stanton, R. V.; Cheng, A. L.; Vincent, J. J.; Crowley, M.; Tsui, V.; Gohlke, H.; Radmer, R. J.; Duan, Y.; Pitera, J.; Massova, I.; Seibel, G. L.; Singh, U. C.; Weiner, P. K.; Kollman, P. A. *AMBER 7*; University of California: San Francisco, 2002.

(23) Cornell, W. D.; Cieplak, P.; Layly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179–5197.

(24) Warshel, A. *Nature* **1976**, *260*, 679.

(25) Warshel, A.; Chu, Z. T.; Hwang, J.-K. *Chem. Phys.* **1991**, *158*, 303–314.

(26) Warshel, A.; Chu, Z. T. *J. Phys. Chem. B* **2001**, *105*, 9857–9871.

(27) Ren, L.; Martin, C. H.; Wise, K. J.; Gillespie, N. B.; Luecke, H.; Lanyi, J.; Spudich, J. L.; Birge, R. R. *Biochemistry* **2001**, *40*, 13906–13914.

(28) Rajamani, R.; Gao, J. *J. Comput. Chem.* **2002**, *23*, 96–105.

(29) Strambi, A.; Coto, P. B.; Ferré, N.; Olivucci, M. *Theor. Chem. Acc.* **2007**, *118*, 185–191.

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

(31) Andersson, K.; Aquilante, F.; Barysz, M.; Bednarz, E.; Bernhardsson, A.; Blomberg, M. R. A.; Carissan, Y.; Cooper, D. L.; Cossi, M.; Devarajan, A.; Vico, L. D.; Ferré, N.; Fülscher, M. P.; Gaenko, A.; Gagliardi, L.; Ghigo, G.; Graaf, C. d.; Hess, B. A.; Hagberg, D.; Holt, A.; Karlström, G.; Krogh, J. W. R.; Lindh, R.; Malmqvist, P.-A.; Nakajima, T.; Neogrády, P.; Olsen, J.; Pedersen, T. B.; Raab, J.; Reiher, M.; Roos, B. O.; Ryde, U.; Schimmelpfennig, B.; Schütz, M.; Sadlej, A. J.; Schütz, M.; Seijo, L.; Serrano-Andrés, L.; Siegbahn, P. E. M.; Stålring, J.; Thorsteinsson, T.; Veryazov, V.; Widmark, P.-O. Molcas, Version 6.2; University of Lund: Lund, Sweden, 2003.

(32) Besler, B.; Merz, K.; Kollman, P. *J. Comput. Chem.* **1985**, *11*, 431–439.

(33) Olivucci, M.; Sinicropi, A. *Computational Photochemistry*; Olivucci, M., Ed.; Elsevier: Amsterdam, 2005; pp 1−33.

(34) Blatz, P. E.; Liebman, P. *Exp. Eye Res.* **1973**, *17*, 573–580.

(35) Kakitani, H.; Kakitani, T.; Rodman, H.; Honig, B. *Photchem. Photobiol.* **1985**, *41*, 471–479.

(36) Neitz, M.; Neitz, J.; Jacobs, G. H. *Science* **1991**, *252*, 971–974.

(37) Frutos, L. M.; Andruniów, T.; Santoro, F.; Ferré, N.; Olivucci, M. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 7764–7769.

(38) Serrano-Andres, L.; Borin, A. C. *Chem. Phys.* **2000**, *262*, 267–283.

(39) Morton, R. A.; Pitt, G. A. J. *Biochemistry J.* **1955**, *59*, 128–134.

(40) Arnaboldi, M.; Motto, M. G.; Tsujimoto, K.; Balogh-Nair, V.; Nakanishi, K. *J. Am. Chem. Soc.* **1979**, *101*, 7082–7084.

(41) Liu, T.; Callis, P. R.; Hesp, B. H.; deGroot, M.; Buma, W. J.; Broos, J. *J. Am. Chem. Soc.* **2005**, *127*, 4104–4113.

(42) Olivucci, M.; Lami, A.; Santoro, F. *Angew. Chem., Int. Ed.* **2005**, *12*, 5118–5121.

JP9080993