JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

Role of the Molecular Environment in Flavoprotein Color and Redox Tuning: QM Cluster versus QM/MM Modeling

Anikó Udvarhelyi,[†] Massimo Olivucci,[‡,§,#] and Tatiana Domratcheva*[,†]

[†]Department of Biomolecular Mechanisms, Max Planck Institute for Medical Research, Jahnstrasse 29, 69120 Heidelberg, Germany

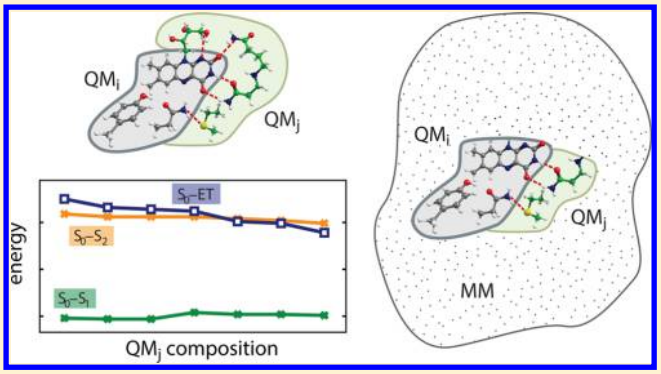
[‡]Dipartimento di Biotecnologie, Chimica e Farmacia, via A. Moro 2, Universitá di Siena, I-53100 Siena, Italy

[§]Chemistry Department, Overman Hall, Bowling Green State University, Bowling Green, Ohio 67200, United States

[#]Institut de Physique et de Chimie des Materiaux de Strasbourg, Université de Strasbourg, Batiment 69, 23 Rue du Loess, 67200 Strasbourg, France

Supporting Information

ABSTRACT: We investigate the origin of the excitation energy shifts induced by the apoprotein in the active site of the bacterial photoreceptor BLUF (Blue Light sensor Using Flavin adenine dinucleotide). In order to compute the vertical excitation energies of three low-lying electronic states, including two $\pi-\pi^*$ states of flavin (S_1 and S_2) and a $\pi-\pi^*$ tyrosine-flavin electron-transfer state (ET), with respect to the energy of the closed-shell ground state (S_0), we prepared alternative quantum mechanical (QM) cluster and quantum mechanics/molecular mechanics (QM/MM) models. We found that the excitation energies computed with both types of models correlate with the magnitude of the charge transfer character of the excitation. Accordingly, we conclude that the small charge transfer character of the light absorbing S_0-S_1 transition and the substantial charge transfer character of the nonabsorbing but redox active S_0-ET transition explain the small color changes but substantial redox tuning in BLUF and also in other flavoproteins. Further analysis showed that redox tuning is governed by the electrostatic interaction in the QM/MM model and transfer of charge between the active site and its environment in the QM cluster. Moreover, the wave function polarization of the QM subsystem by the MM subsystem influences the magnitude of the charge transfer, resulting in the QM/MM and QM excitation energies that are not entirely consistent.



1. INTRODUCTION

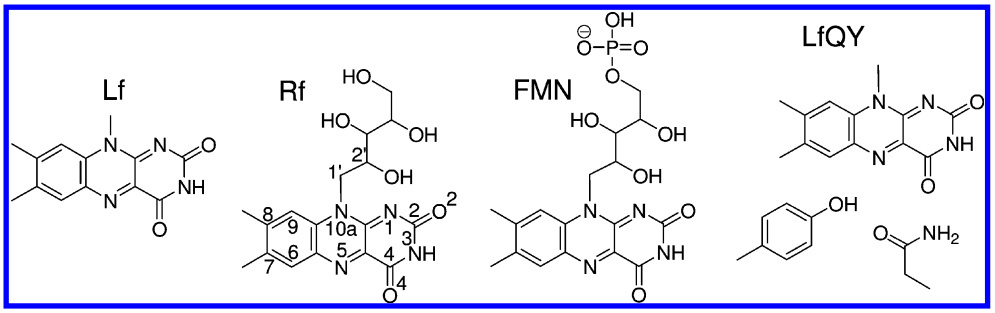
Photosensory flavoproteins activate phototactic responses, circadian adaptation, and even neuronal activity.[1−3] In recent years, these proteins were used in numerous optogenetic applications and as prototypes for engineering fluorescent probes and biosensors.[4−6] The sensory function of these proteins and the fluorescence lifetime depend on the flavin cofactor excitation energies and redox potential. Hence, an understanding of how the interaction between cofactor and apoprotein occurs at the electronic structure level is critical for the rational development of flavoprotein engineering technologies.

Oxidized flavins (see Lf, Rf, and FMN in Scheme 1), representing the chromophore moiety in photosensory flavoproteins, exhibit absorption maxima at 440 nm (2.82 eV) and 370 nm (3.35 eV), corresponding to the S_0-S_1 and S_0-S_2 vertical excitation energies, respectively.[7−11] The S_0-S_1 energy defines the yellow color of oxidized flavoproteins, and usually, the flavin−apoprotein interactions induce only small changes of this energy. Besides the cofactor S_1 and S_2 states, flavoproteins feature low-lying states associated with the transfer of an electron from a protein redox-active residue (the electron donor) to flavin, which we call the intermolecular electron-transfer (ET) states. All characterized photosensory flavoproteins feature ET states involving tyrosine, tryptophan, or cysteine residues as electron donors to flavin. One- and two-electron reduction, which is central for the flavin biological role of a versatile redox cofactor,[12] is also attributed to intermolecular ET states. The S_0-ET excitation energy governing flavin reduction is rather sensitive to flavin-apoprotein interactions. In the excitation spectrum, however, the S_0-ET transition does not carry sufficient oscillator strength (i.e., ET is a "dark" state in the spectrum), and thus, its energy cannot be easily measured in contrast to the S_0-S_1 and S_0-S_2 excitation energies. As a consequence, in spite of its functional importance, there is presently little alternative to the use of quantum chemical methods for the determination of the S_0-ET excitation energy value. For this reason, and also because the number of quantum chemistry computations of the

Received: February 28, 2015
Published: July 7, 2015

DOI: 10.1021/acs.jctc.5b00197
J. Chem. Theory Comput. 2015, 11, 3878−3894

Scheme 1. Chemical Structure of Lumiflavin (Lf), Riboflavin (Rf, with conventional numbering of selected atoms), Flavin Mononucleotide (FMN), and the BLUF Photoactive Triad Lumiflavin-Glutamine−Tyrosine (LfQY)



flavoprotein excited states is constantly growing,[13−22] a critical assessment of the performance of the methods and models available for these computations is highly desirable.

So far computational studies of flavoproteins have been predominantly focusing on the excited states of oxidized flavin cofactors (reviewed in ref 23). Various methods, including TD-DFT methods, proved to be sufficiently accurate in computing the oxidized flavin UV−vis absorption spectrum.[14,24−27] The S_0-S_1 and S_0-S_2 bands were assigned to the $\pi-\pi^*$ HOMO−LUMO and (HOMO−1)−LUMO excitations,[24] respectively. In a polar environment, the band maxima shift to the longer wavelength because S_1 and, even more, S_2 have a larger dipole moment (i.e., are more polar) compared to S_0.[24,25] The S_0-ET transition, described by a low spatial overlap[28] of the HOMO of the electron-donor and the flavin LUMO, is more challenging to compute. The self-interaction error of TD-DFT methods leads to an underestimated S_0-ET energy value.[15] Although the CAM-B3LYP functional[29] and the combined TDA-CAM-B3LYP approach[30] overcome the long-range charge transfer problem (yet, their performance has not been tested for flavoproteins to the best of our knowledge), high-level quantum chemistry methods that are inherently free from the charge transfer problem represent a superior choice.

Ideally, accurate correlated quantum chemical methods are also necessary to account for interactions of the active site with the surrounding protein. Due to sensitivity of the S_0-ET excitation energy to the cofactor−apoprotein interactions, including in the molecular model at least part of the environment surrounding the electron donor and acceptor is essential. Such models are typically constructed via two common approaches: quantum mechanical (QM) cluster approach[31] and quantum mechanics/molecular mechanics (QM/MM) method with electrostatic embedding.[32−37] It is presently not clear if the alternative QM cluster and QM/MM models of a flavoprotein yield consistent results. The present contribution attempts to answer this question by focusing on a specific photosensory flavoprotein called BLUF.

The bacterial photoreceptor domain BLUF (sensor of Blue Light Using FAD)[38] is a small (ca. 120 residues long) globular sensory domain found in complex multidomain proteins or as a single-domain protein. Photoexcitation of BLUF results in photoinduced electron transfer from a tyrosine residue to the FAD cofactor.[8,14,16,26] A conserved glutamine residue mediates the interactions between the flavin electron acceptor and the nearby tyrosine electron donor.[39−42] Thus, in order to computationally model BLUF, a complex of lumiflavin, glutamine, and tyrosine, termed photoactive triad LfQY (Scheme 1), constitutes the minimum-size molecular model. In our previous studies,[13,14,21,43] we established an approach to compute the excitation energies of flavoproteins with the CASSCF method[44] in conjunction with second-order perturbation theory XMCQDPT2.[45] In the present study, we expand our previous calculations by characterizing the interaction between the LfQY triad and its molecular environment in BLUF. More specifically, we are interested in elucidating the interaction energy contributions responsible for the protein color and redox tuning.

In this work, we characterize alternative QM cluster and QM/MM models, both describing the interaction between the LfQY triad and its environment in the PixD BLUF protein from a cyanobacterium.[46] For the QM calculations, we employ the established CASSCF protocol[14] in conjunction with either the XMCQDPT2 method (for the QM cluster) or MS-CASPT2[47] method (for the QM subsystem of the QM/MM model). The choice of the PT2 method was imposed by the available software packages: Firefly[48] for the XMCQDPT2 calculations of the QM cluster and MOLCAS[49] for the CASPT2 and MS-CASPT2 calculations of the QM/MM model. Our QM cluster corresponds to a large fragment of the flavin-binding pocket including the LfQY triad. Notice that the computational efficiency of the XMCQDPT2 code in Firefly enabled us to compute a QM cluster, which is larger in size compared to the QM subsystem of our QM/MM model. The QM/MM model treats the LfQY triad as the QM subsystem and the entire BLUF protein surrounded by a water solvation shell as the MM subsystem. The coupling between the QM and MM subsystems is treated by means of the "electrostatic potential fitting" (ESPF) operator[50] as detailed in Subsection 3.6. In the geometrically optimized QM cluster and QM/MM models, we computed the S_0-S_1, S_0-S_2, and S_0-ET vertical excitation energies. We explain the differences between the QM-cluster and QM/MM results by analyzing the following key factors: (i) geometry optimization constraints, (ii) performance of the MS-CASPT2 and XMCQDPT2 methods, (iii) amount of charge transfer within the photoactive LfQY triad upon the excitation, and (iv) electrostatic interaction and wave function polarization interaction and transfer of charge between the LfQY triad and its environment.

2. COMPUTATIONAL DETAILS

2.1. QM Cluster Model. Using the PDB coordinates of the PixD BLUF protein (PDB code 2HFN, chain A),[41] the starting geometry of the RfQY-MSNNL QM cluster was prepared. The QM cluster consists of the riboflavin (Rf) portion of the flavin mononucleotide (FMN) cofactor and the residues Q53, Y11, M96, S31, N34, N35, and L44 (QY-MSNNL). We included the M96 and L44 side chains and the S31, N34, and N35 side chains together with selected backbone atoms. To mimic the

DOI: 10.1021/acs.jctc.5b00197
J. Chem. Theory Comput. 2015, 11, 3878−3894

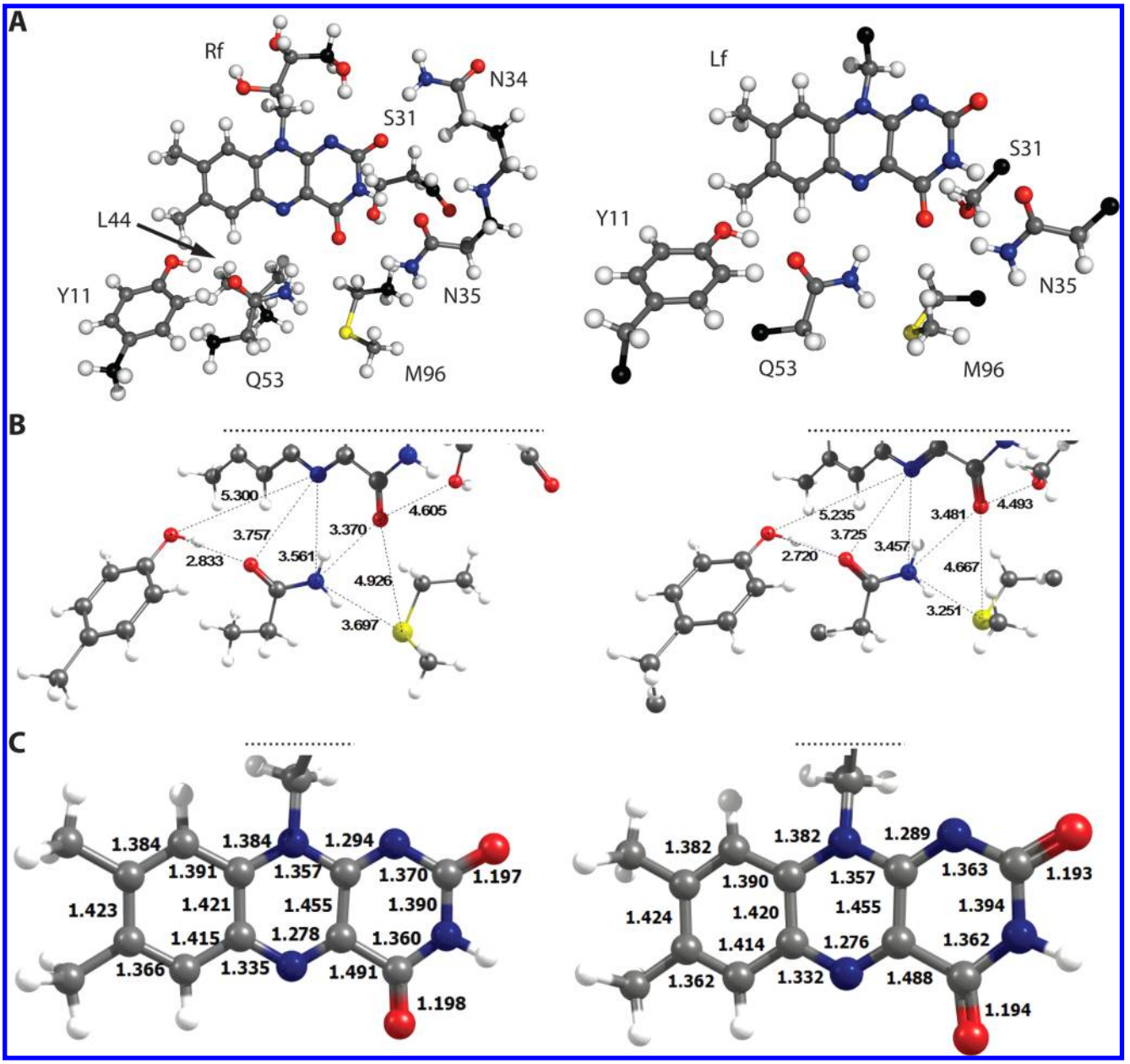


Figure 1. Optimized geometry of the BLUF active site in the QM cluster and QM/MM model. (A) RfQY-MSNNL cluster optimized with the constrained positions of the black-colored carbon atoms and of the S31 hydroxyl oxygen (left panel). The largest QM subsystem employed in the QM/MM calculations is shown with black-colored boundary MM carbon atoms (right panel). (B) Selected inter-fragment equilibrium distances (Å) in the QM cluster (left) and QM/MM model (right). The N35 side chain is not shown for clarity. (C) Selected equilibrium bond distances (Å) in the isoalloxazine moiety in the QM cluster (left) and QM/MM model (right).

"stiffness" of the active site in the protein, the distances between all pairs of selected terminal atoms (the black-colored carbon atoms of each fragment indicated in Figure 1A and the oxygen atom of S31) were fixed to the corresponding value in the crystal structure. The geometry optimization was performed in the electronic ground state with the CASSCF(6,4)4 method (six-electrons-in-four-MOs performing energy averaging over four states with equal weights) in the delocalized internal coordinates[51] with the described distance constraints. The HOMO−1, HOMO, and LUMO of flavin and the HOMO of tyrosine (of the Hartree−Fock reference wave function) were included in the active space. At the optimized geometry, the excitation spectrum was computed with the XMCQDPT2 method with the zeroth order CASSCF(6,4)4 wave function and with the intruder-state avoidance denominator shift[52] of 0.02 au.

To analyze the role of specific intermolecular interactions, we prepared a set of QM clusters by successively excluding selected molecular fragments from the energy minimized RfQY-MSNNL cluster. The influence of the interaction between the isoalloxazine chromophore and the ribityl chain was estimated by substituting riboflavin (Rf) with lumiflavin (Lf). To obtain the QM clusters with the Lf chromophore, the atoms composing the Rf ribityl fragment were removed and one H atom was placed along the C1′−C2′ bond (Scheme 1) at a distance of 1.091 Å from the C1′ atom. All computed reduced QM clusters are listed in Section S1 of the Supporting Information. For each reduced cluster, the XMCQDPT2

DOI: 10.1021/acs.jctc.5b00197
J. Chem. Theory Comput. 2015, 11, 3878−3894

excitation energies were computed without reoptimizing its geometry.

The XMCQDPT2 calculations of isolated Lf and Rf molecules were carried out with the CASSCF(4,3)3 zeroth order wave function (four-electrons-in-three-MOs performing energy averaging over three states with equal weights) with the HOMO−1, HOMO, and LUMO in the active space. To demonstrate the effect of the active space expansion on the XMCQDPT2 excitation energies, the computations of the Lf molecule and LfQY complex were conducted with several active spaces constructed by selecting the frontier π and π^* MOs of flavin and tyrosine according to their Hartree−Fock orbital energies. In all cases, we performed state-averaged CASSCF calculations with equal weights for four states (LfYQ) or three states (Lf). For the purpose of comparison, the excitation energies of the Lf molecule and LfYQ complex were computed at the geometry optimized with the B3LYP density functional method. When comparing the results of the cluster models of varying size, the basis set superposition error (BSSE) was estimated following the counterpoise correction scheme.[53] In addition, the effect of size inconsistency was evaluated by comparing the XMCQDPT2 excitation energies to those computed with the core-consistent XMCQDPT2′ method.[45]

To analyze the electrostatic charge distribution, we computed molecular electrostatic potentials (MEPs) and Mulliken atomic charges for each electronic state with the zeroth order QDPT density. The MEPs computed on the "fine" grid (approximately 100 × 100 × 100 points) as defined in the Firefly program, were visualized with the Chemcraft program.[54] The spheres used for the visualization were scaled by a factor of 1.5 to improve the visibility of the differences in the figures presented below. All calculations of the cluster model were performed with the Firefly quantum chemistry package,[48] which is partially based on the GAMESS US source code.[55] The cc-pVDZ basis set (obtained via the Basis Set Exchange software and the EMSL Basis Set Library[56,57]) was used throughout.

2.2. QM/MM Model. We prepared the initial coordinates of the QM/MM model with the help of the GROMACS 4.5 package.[58] The model was prepared starting from the same BLUF protein structure as in the case of the QM cluster (PDB code 2HFN, chain A). The hydrogen atoms were added to the PDB model assuming standard protonation states (arginines and lysines protonated, aspartates and glutamates deprotonated, histidines neutral Nδ-H form) except for histidine H75 assigned to the neutral Nε-H form. A 7.55 nm × 7.55 nm × 7.55 nm rhombic dodecahedron periodic box, containing 9055 water molecules and 7 Na^+ ions, was employed. The entire model was subjected to 500 steps of steepest-descent energy minimization with the GROMOS53A6 force field[59] (because it includes parameters for FMN) and the SPC water model.[60]

For the QM/MM calculations, the above energy-minimized model was reduced in size to include the protein, Na^+ counterions, and a water-shell of 15 Å around the protein (4519 water molecules). The QM subsystem of the QM/MM model contains lumiflavin (Lf) and the Y11 and Q53 side chains, whereas the rest of the protein with the solvent is described by the MM subsystem (indicated below as LfQY/MM model). The QM subsystem interacts electrostatically with the MM environment by means of the ESPF operator[50] (see Subsection 3.6 for further details). The computations were carried out employing the QM/MM coupling scheme between the MOLCAS[49] and Tinker[61] programs. Such a methodology has been widely employed to compute the excitation energies of biological photoreceptors,[62,63] as it allows the use of the multiconfigurational CASPT2//CASSCF or MS-CASPT2//CASSCF method (excitation energy computed with CASPT2 at the geometry optimized with CASSCF) in treating the QM subsystem.

The AMBER force field[64] and the TIP3P water model[65] were used to describe the MM subsystem. For the FMN cofactor with the protonated phosphate (charge −1), the RESP charges were derived from the HF/6-31G* calculation using the antechamber package.[66] The atom types and van der Waals parameters of FMN were assigned by analogy to the AMBER atom types of amino acids and nucleobases. The hydrogen link atom (LA) scheme[67] was used at the QM/MM boundary. The frontier was placed between nonpolar sp^3 carbon atoms C1′−C2′ in the FMN cofactor, Cα−Cβ in Y11, and Cβ−Cγ in Q53. To reduce overpolarization of the QM subsystem, the charge of the boundary MM carbon atom was set to zero and equally redistributed among the three neighboring MM atoms. In addition, the small residual charges of the MM part of the partitioned side chains and FMN were also equally redistributed among the mentioned three MM atoms.

We optimized the geometry of the LfQY/MM model in the ground state with the CASSCF(6,4)4 method, with "micro-iterations[68] on" to relax the geometry of the MM subsystem within 4 Å of any QM atom, whereas the remaining MM atoms we kept frozen. At the equilibrium geometry, we computed the excitation energies with the CASPT2 and MS-CASPT2 methods with an imaginary shift[69] of 0.2 au and without applying the IPEA shift.[70] In all calculations with MOLCAS we used the DZV(P) (d-functions on the heavy atoms) basis set.

Modifications of the LfQY/MM model were used to investigate the effect of the molecular environment on the vertical excitation energy spectrum. The model variants were constructed by repartitioning the QM and MM subsystems such that the QM subsystem includes, in addition to the LfQY triad, the side chains S31, N35, and/or M96. The new QM/MM boundary was introduced between the Cα and Cβ atoms of S31 and N35 and between the Cβ and Cγ atoms of M96, and the same charge-neutralization scheme as in the LfQY/MM model was used. The geometry of the model variants was kept as obtained after the initial geometry optimization of the LfQY/MM model (that is, without geometry reoptimization); hence, the residues added to the extended QM subsystems have the geometries optimized according to the MM force-field parameters. For the LfQY triad and for all extended QM subsystems, we conducted the excited-state calculations with the CASPT2, MS-CASPT2, and XMCQDPT2 methods. Excitation energies of the Lf molecule and the Lf/MM model (Lf in the QM subsystem and the phosphorylated ribityl chain with the solvated BLUF apoprotein in the MM subsystem) were computed with the CASSCF(4,3)3 zeroth order wave function.

In addition, in order to investigate electrostatic interactions, we computed excitation energies for the modified LfQY/MM^{off} and LfQY/$MM^{+1/-1}$ models. The LfQY/MM^{off} models were obtained by zeroing the MM charges of the S31, N35, or M96 residues in the initially optimized LfQY/MM model with the same treatment of the QM/MM boundary as described above. In the LfQY/$MM^{+1/-1}$ models, an extra +1 or −1 charge was placed on molecular fragments in the vicinity of the electron donor and the electron acceptor. These +1/−1 MM-charge modifications were carried out at the initially obtained LfQY/

DOI: 10.1021/acs.jctc.5b00197
J. Chem. Theory Comput. 2015, 11, 3878−3894

Table 1. Dependence of the XMCQDPT2 Vertical Excitation Energies (eV) of the BLUF Active Site on the Choice of the Zeroth Order CASSCF Wave Function[a]

	lumiflavin			
	(4,3)3	(8,8)3	(12,12)3	(14,13)3
S_0-S_1	**3.09** [0.317]	2.90 [0.219]	2.96 [0.224]	2.90 [0.197]
S_0-S_2	**4.23** [0.287]	4.11 [0.149]	4.28 [0.150]	4.20 [0.148]
	BLUF photoactive triad LfQY			
	(6,4)4	(14,12)4	(14,13)4	(14,13)4
Lf[b]	**(4,3)**	(8,7)	(12,11)	(8,8)
Y[c]	**(2,1)**	(6,5)	(2,2)	(6,5)
S_0-S_1	**3.01** [0.307]	3.26 [0.257]	2.88 [0.192]	3.60 [0.222]
S_0-S_2	**4.04** [0.282]	4.17 [0.283]	4.21 [0.184]	4.27 [0.278]
S_0-ET	**4.44** [0.005]	4.53 [0.004]		4.53 [0.004]
S_0-S_3[d]			5.06 [0.083]	

[a]Transition dipole moments (au) are indicated in square brackets. The CASSCF zeroth order wavefunction is specified by the size of the active space and number of states computed, (# electrons, # MOs, # states). The active spaces and corresponding energies used for further analysis are indicated in bold. [b]Subspace of CAS localized on lumiflavin (Lf). [c]Subspace of CAS localized on tyrosine (Y). [d]S_0-S_3 corresponds to the $\pi-\pi^*$ transition in Lf.

MM geometry but with a modified QM/MM boundary, in which only the charge of the boundary MM carbon atom was redistributed among the three neighboring MM atoms, but the small charge on the MM portions of FMN, Y11, and Q53 was not redistributed. In the LfQY/MM model, this boundary modification practically does not change the S_0-S_1 and S_0-S_2 energies but decreases the S_0-ET energy by 0.11 eV.

For the QM subsystems with and without embedding in the MM subsystem, the MEPs were computed from the CASSCF electron densities on a grid that was similar to the "coarse" grid (approximately 40 × 40 × 40 grid points) defined in the Firefly program. The coordinates of the grid points were explicitly specified in the MOLCAS input file. For the MEP visualization with Chemcraft, cube files were prepared from the MOLCAS outputs with a custom-written Perl script. Atomic Mulliken and ESPF charges were computed with the CASSCF and CASPT2 methods as implemented in MOLCAS 7.6.

3. RESULTS AND DISCUSSION

3.1. Choice of the CASSCF Active Space. We computed vertical excitation energies in molecular complexes employing the zeroth order CASSCF wave function and perturbation theory energy correction to the second order: XMCQDPT2, MS-CASPT2, and CASPT2. In the active space, we include MOs essential to describe the static correlation of the four states of interest, i.e., the closed-shell state S_0, locally excited states of flavin S_1 and S_2, and the ET state of intermolecular electron transfer from tyrosine to flavin. The multireference perturbation theory calculations recover the dynamical correlation for all computed states. The active space to account for static correlation is rather small, and the respective CASSCF excitation energies are rather high, whereas the energies obtained with the perturbation theory treatment are in good agreement with the experimental values and also with the results of other methods, which do not depend on the choice of the active space, as demonstrated in our previous study.[14] In this study, we analyze vertical excitation energies in the BLUF active site computed with cluster and QM/MM models to which in further description we refer as the S_0-S_1, S_0-S_2, and S_0-ET energies. We included the HOMO−1, HOMO, and LUMO of the flavin and the HOMO of tyrosine in the (6,4) active space because the three excited states in question are dominated by single-electron excitations. Examples of computations with reduced active spaces have been also reported for the green fluorescent protein (GFP) chromophore,[71] malonaldehyde,[72] and butadiene.[73]

An alternative and more common approach is to select an active space as large as possible, focusing on the $\pi-\pi^*$ subsystem.[74] In practice, as the number of configurations increases very rapidly upon active-space expansion, active spaces larger than (14,14) are seldom computed. As the $\pi-\pi^*$ subsystem of a flavin and tyrosine complex exceeds the computationally affordable size, only some of these orbitals can be included in the active space. We considered several such extended active spaces and compared the resulting XMCQDPT2 excitation energies with those obtained with the (6,4) active space (Table 1). In addition, we demonstrate how the active-space expansion influences the XMCQDPT2 excitation energies of the isolated Lf molecule. As the results demonstrate that the gradual active-space extension does not successively improve the XMCQDPT2 excitation energies, we use the (6,4) active space for our further analysis.

In our study, we employed CASSCF calculations for geometry optimization. The small-active-space CASSCF geometry of the closed-shell ground state is similar to the one obtained with the Hartree−Fock method and is inferior to the commonly used DFT-optimized geometry. The excitation energies obtained at the CASSCF geometry are somewhat overestimated. In the Lf molecule, the single-point XMCQDPT2-CASSCF(4,3)3 S_0-S_1 and S_0-S_2 energies computed at the B3LYP geometry are 3.05 and 4.13 eV, respectively, which is slightly lower than the energies computed at the CASSCF-optimized geometry presented in Table 1. In the photoactive triad, the XMCQDPT2-CASSCF(6,4)4 excitation spectrum at the B3LYP geometry is 2.89 eV (S_0-S_1), 3.91 eV (S_0-S_2), and 4.10 eV (S_0-ET). The increased deviation of the S_0-S_2 and S_0-ET energies when the results of the two geometries are compared is explained by the noticeable shortening of the hydrogen bond distances in the B3LYP geometry. Yet, the multireference CASSCF method presents a unique tool for optimizing geometry in a degenerated ground state, which is of interest for our studies in progress. Therefore, we conducted our analysis of intermolecular-interaction energies using the CASSCF-optimized geometries. It is noteworthy, that the excited state stabilization at the B3LYP geometry increases with an increase

DOI: 10.1021/acs.jctc.5b00197
J. Chem. Theory Comput. 2015, 11, 3878−3894

in the excitation energy. The analysis presented in the following subsections documents that such an effect is related to the increased amount of charge transfer in the corresponding electronic transitions.

3.2. Optimized Geometry and Protein Constraints. The optimized geometries of the flavin-binding site obtained with the QM cluster and with the QM/MM model are somewhat different (Figure 1). The QM/MM active-site geometry features intermolecular distances close to those of the crystal structure. In contrast, the equilibrium QM cluster geometry is significantly displaced from the crystal structure, and most of the distances characterizing the interfragment interactions increase in the optimized QM cluster as compared to the QM/MM model (Figure 1B). For instance, the S31 residue slightly moves away from its initial position, thus disrupting the hydrogen bond between the S31 backbone carbonyl and the N35 backbone nitrogen. The displacement of the M96 side chain, in contrast, enables the formation of a hydrogen bond between flavin C4═O4 and NH_2 of Q53 reducing the corresponding distances. The role of the M96 side chain in modulating the flavin-Q53 interaction has been noticed in a previous computational study.[75] At the same time, the optimized geometries of the isoalloxazine system (Figure 1C) are similar in the QM cluster and QM/MM model.

Our observation that the QM cluster yields a less compact geometry of the active site than the QM/MM model is, in fact, not surprising. Choosing a more conservative coordinate locking scheme and/or enlarging the size of the molecular cluster (as for example in our previous calculations in ref 27) should decrease the structural relaxation and minimize the geometry deviations from the crystal structure. In the following, we demonstrate how the geometry differences between the QM cluster and QM/MM model, i.e., the elongated hydrogen bonds and increased donor−acceptor distance, influence the excitation spectrum of the flavoprotein active site.

3.3. Excitation Energies Computed with the QM Cluster. In our description, we specify the cluster composition by indicating the photoactive triad FlQY, followed by the sequence of the residues included in the cluster using the standard one-letter amino-acid residue notation. In this notation, our largest computed cluster model is RfQY-MSNNL (Figure 1). The S_0−S_1 and S_0−S_2 energies computed with the RfQY-MSNNL cluster at the XMCQDPT2 level of theory are 3.07 and 3.82 eV, respectively, and the S_0−ET energy is 4.06 eV. The comparison of these energies with the results of the reduced-size clusters (in which some of the molecular fragments were removed without changing the geometry of the retained fragments) demonstrates the effect of intermolecular interactions, described at the CASSCF or XMCQDPT2 level of theory (Figure 2). In the figure, the cluster models are ordered according to the decreasing S_0−ET energy values. In the given order, the small changes of the S_0−S_1 energy are nonmonotonous while the larger changes of the S_0−S_2 energy follow the prominent decrease in the S_0−ET energy. The latter trend indicates that a BLUF protein with a red-shifted S_0−S_2 absorbance maximum may undergo accelerated photoinduced electron transfer. Interactions with residues N34, S31, and M96 decrease the S_0−S_2 and S_0−ET energies, in contrast to interaction with the N35 residue. The CASSCF excitation energy values are significantly higher than the corresponding XMCQDPT2 ones; however, the energy shifts from one cluster model to the other computed with both methods are rather consistent. An exception is the S_0−S_1

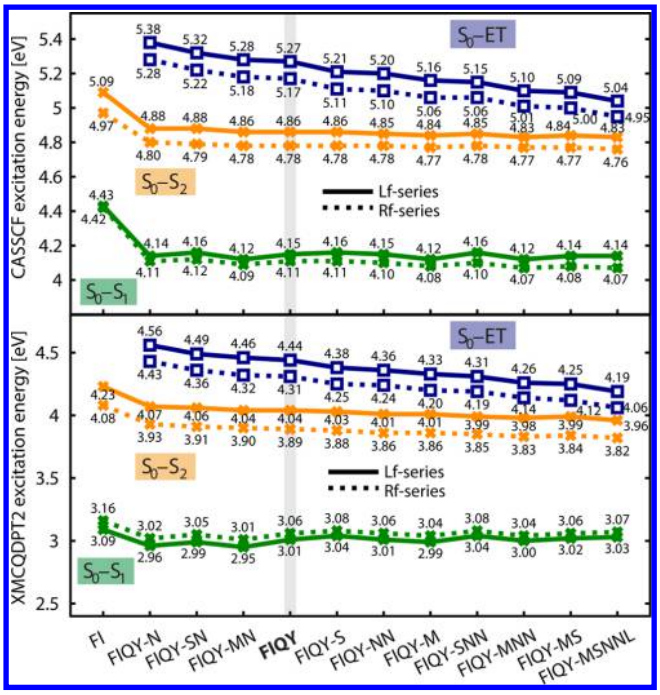


Figure 2. CASSCF and XMCQPDT2 excitation energies computed for the QM clusters. Fl (flavin) indicates either Lf (lumiflavin) or Rf (riboflavin). FlQY indicates the photoactive triad. The other letters indicate the residues included in the cluster according to the standard one-letter amino-acid residue notation. One asparagine "N" in the model designation corresponds to N35, whereas two asparagines "NN" correspond to N34 and N35.

energy change upon removing the ribityl chain of the flavin cofactor (the CASSCF S_0−S_1 energy increases whereas the XMCQDPT2 S_0−S_1 energy decreases in the Lf-series compared to the Rf-series).

To further validate the comparison of the models varying in molecular size, we computed the BSSE and evaluated the effect of the size inconsistency of the XMCQDPT2 method. The BSSE is largest for the ET state in comparison to the other states as presented in Section S2 of the Supporting Information. The large stabilization of the ET state in the complexes results from its charge-separated character (flavin anion-radical and tyrosine cation-radical). The largest BSSE correction decreases the S_0−ET energy in the RfQY complex from 4.31 to 4.26 eV when the ghost orbitals of the SNNML environment are added. The size-consistent XMCQDPT2′ S_0−ET energies were found to be about 0.02 eV larger as compared to the size-inconsistent XMCQDPT2 ones, independent of the cluster size. In conclusion, neither BSSE nor size inconsistency change the trends revealed in Figure 2. We conclude that the excitation energy changes depending on the QM cluster size are due to intermolecular interactions.

3.4. Charge Transfer and Electrostatic Interaction in the QM Cluster. The three excited states of BLUF display a larger dipole moment (i.e., are more polar) than S_0 (Table 2); therefore, the population of any of these states is characterized by a certain amount of charge transfer. The charge transfer is obvious upon visual inspection of the difference between the molecular electrostatic potentials (MEPs) in the two electronic states (Figure 3A). The MEPs of specific electronic states are presented in Section S4 of the Supporting Information. Upon the S_0−S_1 and S_0−S_2 transitions, about −0.2 and −0.4 au, respectively, is transferred, intramolecularly, from the phenyl to

DOI: 10.1021/acs.jctc.5b00197
J. Chem. Theory Comput. 2015, 11, 3878−3894

Table 2. Dipole Moment (Debye) in the Ground and Excited Electronic States in the QM Clusters Computed with the Zeroth Order QDPT2 Wave Function

QM cluster	S_0	S_1	S_2	ET
LfQY-N	7.98	8.43	13.01	38.96
LfQY	4.92	7.53	12.33	38.47
LfQY-MS	5.75	9.37	12.98	35.12
RfQY-MSNNL	8.11	11.62	15.31	39.91

the uracil moieties of flavin. In the S_0−ET transition, one electron is transferred, intermolecularly, from tyrosine to isoalloxazine. Notably, the larger amount of the transferred charge corresponds to the larger variation of the excitation energy in the QM cluster series: 0.13 eV for S_0−S_1, 0.41 eV for S_0−S_2, and 0.5 eV for S_0−ET. In this subsection, we concentrate on two types of interactions related to the charge redistribution: electrostatic interaction of the FlQY triad atomic charges with the charges of the environment and transfer of electrostatic charge from the triad to its environment. Indeed, the triad environment transfer of charge is evident upon inspection of the MEPs (Figure 3A), especially in the S_0−ET transition. The S_0−ET energy is, among all excitation energies considered here, the most sensitive to intermolecular interactions. Using the XMCQDPT2 results, we establish a correlation between the amount of charge transfer and the S_0−ET energy.

We now consider the electrostatic interaction energy of the RfQY triad with atomic charges q_i, with its environment atomic charges q_j separated by the distances r_{ij} (Figure 3B). The Coulomb energy shift $\Delta\varepsilon_C$ is computed as a difference in the electrostatic interaction energy between the ET state charges and the S_0 state charges, denoted as superscripts in eq 1.

$$\Delta\varepsilon_C = \sum_{i,j}\left(\frac{q_i^{ET}q_j^{ET}}{r_{ij}} - \frac{q_i^{S_0}q_j^{S_0}}{r_{ij}}\right) \quad (1)$$

The Coulomb energy shift can be compared with the energy difference between the S_0−ET energy of the QM cluster containing the environment and the S_0−ET energy of the RfQY triad, which we term the QM energy shift due to the environment, $\Delta\varepsilon_{QM}$ eq 2.

$$\Delta\varepsilon_{QM} = \varepsilon_{QM}[\text{cluster}] - \varepsilon_{QM}[\text{RfQY}] \quad (2)$$

Similarly, throughout this contribution, ε and $\Delta\varepsilon$ denote the excitation energy and the excitation energy shift, respectively. We computed the Coulomb energy shifts in the Rf-series with respect to the energy of the RfQY triad cluster. The QM clusters fall into two groups (Figure 3B): one group with the QM energy shift rather well approximated by the Coulomb

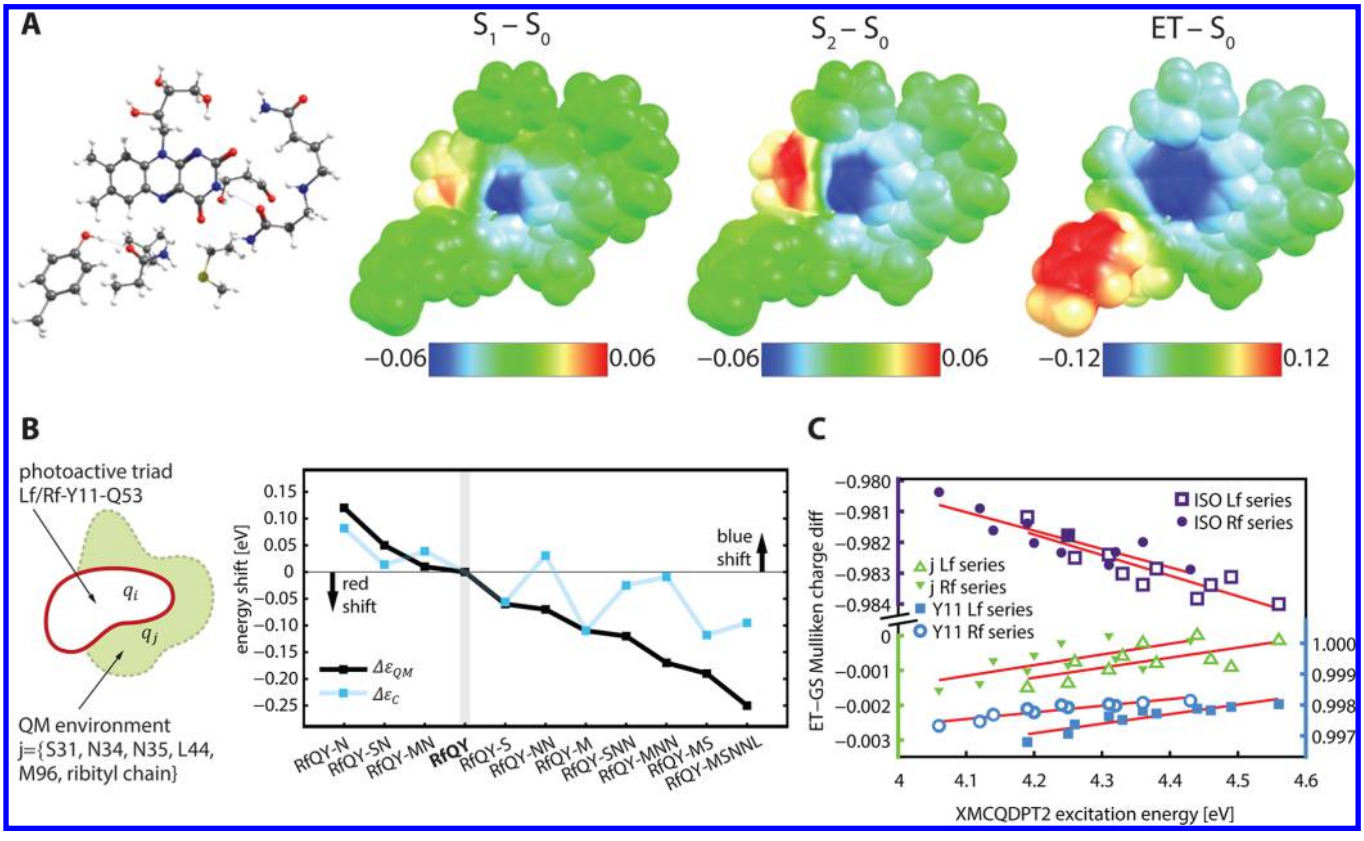


Figure 3. Charge transfer and electrostatic effects. (A) Difference (excited state minus S_0) molecular electrostatic potentials (MEPs) (in au) computed for the RfQY-MSNNL cluster. (B) Schematic representation of the QM cluster partitioning into the FlQY triad with atomic charges q_i and the environment with atomic charges q_j (left panel). Comparison of the QM XMCQDPT2 and Coulomb S_0−ET energy shifts (right panel) as defined in Subsection 3.4. (C) Correlation between the S_0−ET energy and charge transfer magnitude. The difference (ET minus S_0) net-charge was computed on isoalloxazine ISO (dark blue), tyrosine Y11 (light blue), and on the environment j (green). The analysis was separately performed for the QM clusters with riboflavin (Rf series) and lumiflavin (Lf series). The red lines indicate linear regression. In (B) and (C), the Mulliken atomic charges were used.

DOI: 10.1021/acs.jctc.5b00197
J. Chem. Theory Comput. 2015, 11, 3878−3894

energy shift and the other group with the QM energy shift being significantly more negative than the Coulomb energy shift, indicating the presence of nonelectrostatic interactions that stabilize the charge separated ET state and decrease the S_0−ET energy. Overall, the inclusion of more fragments in the QM cluster increases the nonelectrostatic effect. Yet, the Coulomb interactions, as expected, account for a significant part of the excitation energy shift.

The nonelectrostatic interaction of the triad with the environment (given by the deviation between corresponding $\Delta\varepsilon_C$ and $\Delta\varepsilon_{QM}$ values in Figure 3B) becomes more significant in the QM clusters with the reduced S_0−ET energy. The more the S_0−ET energy is shifted to the red, the more the negative charge transfer from tyrosine to flavin deviates from one electron due to the increased leakage of negative charge into the environment of flavin (Figure 3C). In the most red-shifted RfQY-MSNNL cluster, the deviation between $\Delta\varepsilon_C$ and $\Delta\varepsilon_{QM}$ constitutes 0.15 eV, about 0.05 eV of which is due to BSSE and the remaining 0.1 eV might be assigned to the charge transfer stabilization of the charge-separated ET state by the environment. Yet, the magnitude of the leaked charge is only −0.0057 au; however, as previously established,[76−78] transfer of even a small amount of charge due to quantum mechanical interactions stabilizes intermolecular complexes that contain ions. Accordingly, we see significant stabilization of the ET state by the increasing charge leakage if we include the side chains surrounding the flavin in the QM cluster.

Accounting for quantum mechanical interactions of the protein active site with its environment became an important topic in macromolecular simulations in the recent years. In photosensory proteins, the extension of the QM model typically stabilizes charge separated excited states (i.e., yields a red shift of the excitation energy).[79,80] Accounting for quantum mechanical effects either directly[81] or via density embedding methods[82−84] or via introducing charge transfer potentials in the hybrid QM/MM models[77,85−87] is a subject of intensive development. At the same time, the development of efficient computational algorithms for excited state calculations enables computations of increasingly large QM clusters,[88] which ultimately will provide new insights concerning the role of quantum mechanical effects in protein color and redox tuning in the near future.

3.5. Excitation Energies Computed with the QM/MM Model. With the LfQY/MM model treating the QM subsystem at the MS-CASPT2 level of theory, the MM subsystem with the AMBER force field, and the QM/MM coupling with the ESPF operator, we obtained the S_0−S_1 and S_0−S_2 energies of 2.77 and 3.67 eV, respectively, 0.3 and 0.15 eV red-shifted compared to the corresponding values in the largest QM cluster; and the S_0−ET energy of 4.56 eV, which is 0.5 eV blue-shifted with respect to the value in the QM cluster. To infer on the effects causing the differences, we generated a series of QM/MM models, paralleling the QM cluster series, by including the N35, S31, and/or M96 side chains into the QM subsystem (Figure 1A, left panel). For the LfQY triad, as well as for the extended QM subsystems (all having coordinates of the LfQY/MM optimized geometry), we performed CASPT2, MS-CASPT2, and also XMCQDPT2 calculations without including the MM subsystem. By comparing the obtained QM excitation energies of the QM subsystems with the results of the QM clusters, we analyze the effects of the optimized geometry and of the PT2 method. Then we consider the effect of adding the MM environment and the effect of zeroing the MM charges on the N35, S31, and M96 side chains.

The impact of the equilibrium geometry on the computed vertical excitation energies is revealed through the comparison of the CASSCF and XMCQDPT2 values for models of the same composition (QM clusters in Figure 2 and QM subsystems of QM/MM models in Figure 4A). The S_0−S_1 and S_0−S_2 energies computed with the same method at the two geometries are very similar, whereas the S_0−ET energy is 0.3 eV downshifted at the more compact QM/MM geometry. Thus, in agreement with theory, the shorter donor−acceptor distance facilitates intermolecular ET.

The role of the PT2 theory is demonstrated by our calculations of the QM subsystems (Figure 4A). The CASPT2 method gives smaller excitation energies than the XMCQDPT2 method, but the MS-CASPT2 estimates are closer to the XMCQDPT2 ones, yet from 0.1 to 0.3 eV lower. Applying the IPEA shift[70] should increase the CASPT2 and MS-CASPT2 energies toward the XMCQDPT2 values as it was demonstrated in an extensive benchmark study.[89] Another inconsistency between the CASPT2 and XMCQDPT2 energies is found in the models featuring the S_2/ET state crossing upon adding the S31 and M96 side chains into the QM subsystem. The charge redistribution analysis (Figure 4B) indicates that the crossing states mix in the zeroth order CASSCF wave function, causing the spurious CASPT2 results. A similar MS-CASPT2 artifact was previously reported elsewhere[89] (see for instance Figure 2A in ref 89). In contrast to MS-CASPT2, the XMCQDPT2 method is free from the state-crossing artifact;[45] therefore, it gives more reliable S_0−S_2 and S_0−ET energies in the extended QM subsystems.

To investigate the effect of embedding the QM subsystem into the MM subsystem, we compared the QM energies of the isolated QM subsystems (i.e., computed without electrostatic embedding) with their QM/MM counterparts (Figure 5A). At the CASSCF level, we found a downshift of S_0−S_1 and S_0−S_2 energies and a large upshift of S_0−ET energy. At the CASPT2 level, the S_0−S_1 energy is surprisingly upshifted (i.e., the shift changes the sign as compared with the CASSCF result), the S_0−S_2 energy is downshifted, and the S_0−ET energy is significantly upshifted. The decrease in S_0−S_2 energy concomitant with the increase in S_0−ET energy removes the S_2/ET state crossing at both CASSCF and CASPT2 levels of theory and the change in the CASPT2 S_0−S_2 and S_0−ET energies becomes monotonous. Modification of MM charges of the residues in the vicinity of flavin such as zeroing the charges of the S31, N35, or M96 residues in the LfQY/MM^{off} model causes rather small excitation energy variations (Figure 5B) with the largest shift observed for the S_0−ET energy upon zeroing the charges of the S31 residue.

3.6. Charge Transfer, Wave Function Polarization, and Electrostatic ESPF Interaction in the QM/MM Model. The interactions with the MM environment increase the dipole moments of the S_0, S_1, and S_2 states, although to a different extent, and slightly decrease the ET state dipole moment (Table 3). The dipole moment change is caused by polarization of the QM subsystem upon embedding. Consistently, the amount of charge transfer increases for the S_0−S_1 transition, decreases for the S_0−S_2 transition, and almost does not change upon S_0−ET transition (Figure 6). To demonstrate this finding, we report in Section S4 of the Supporting Information the MEPs computed for each of the four states with and without embedding. Polarization of the QM subsystem by the MM

DOI: 10.1021/acs.jctc.5b00197
J. Chem. Theory Comput. 2015, 11, 3878−3894

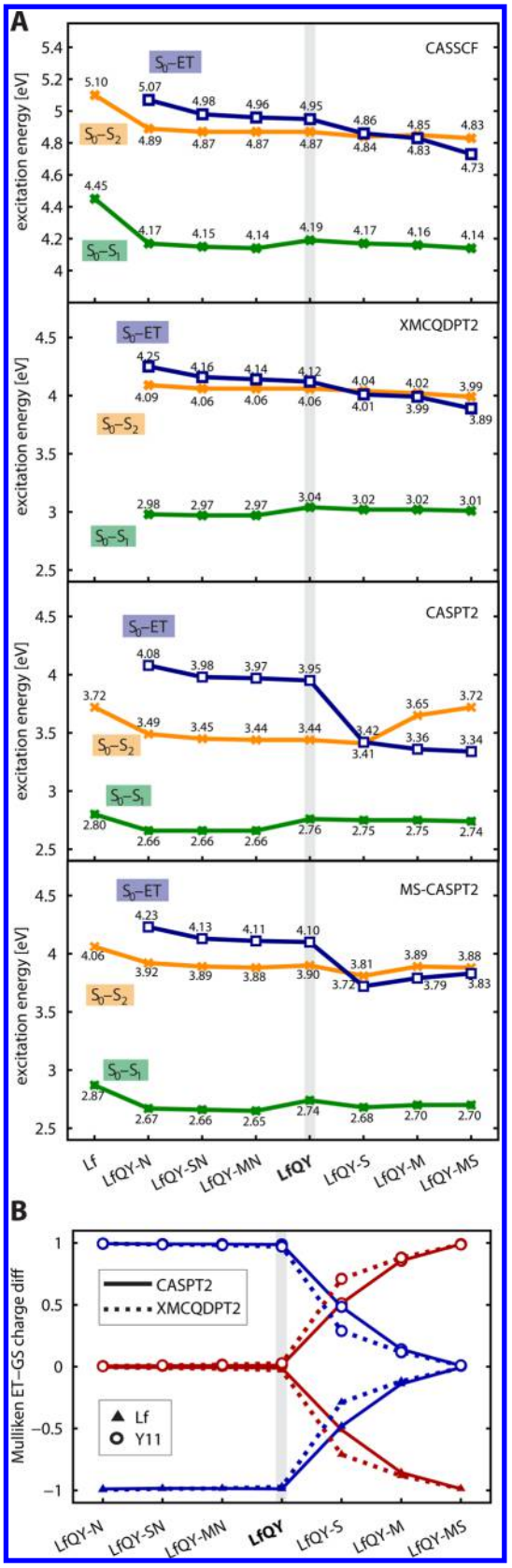


Figure 4. Excitation energies computed for the QM subsystems. (A) CASSCF, XMCQDPT2, CASPT2, and MS-CASPT2 excitation energies. (B) Difference (excited state minus S_0) net charge values were computed using Mulliken charges on the electron donor Y11 and electron acceptor Lf in the second (red line) and third (blue line) excited states to demonstrate the crossing of the S_2 and ET states.

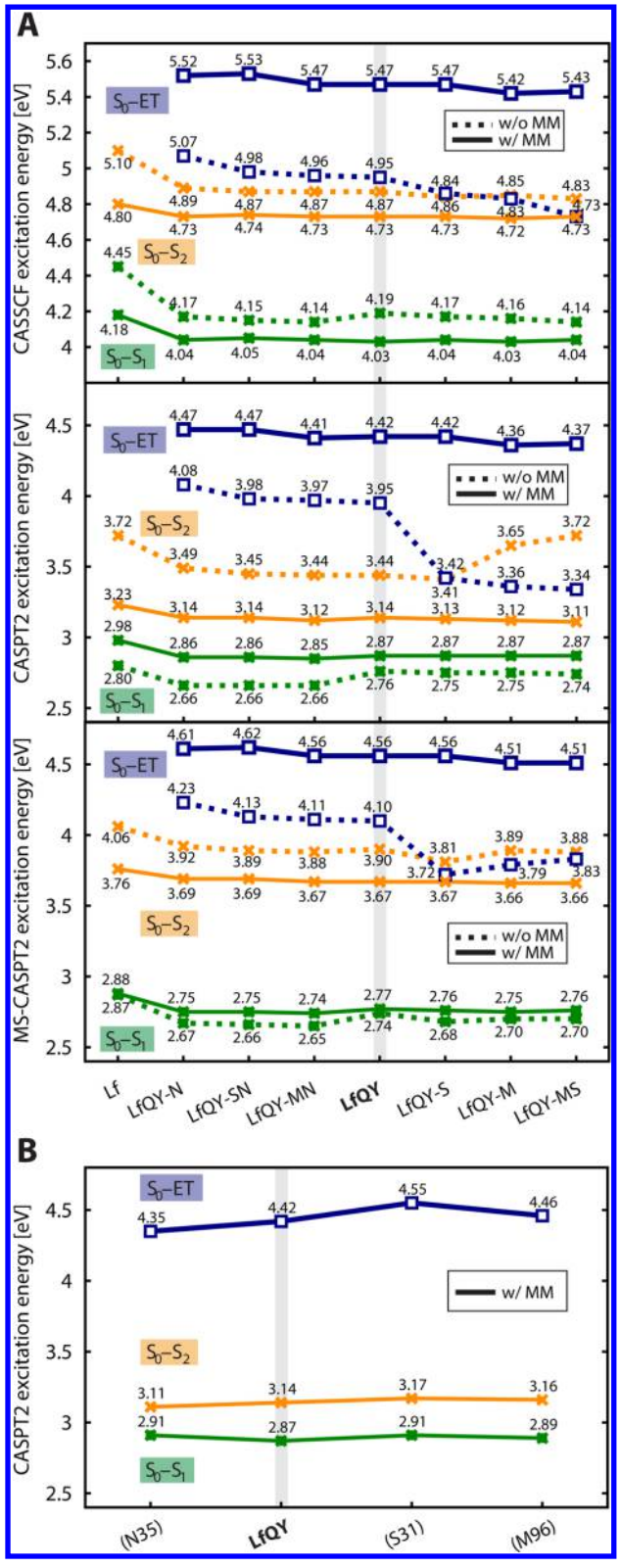


Figure 5. Comparison of the QM/MM and QM excitation energies. (A) Excitation energies of the QM subsystems of various compositions. (B) Effect of the zeroed MM charges in selected residues (indicated in brackets) in the QM/MM^{off} models. In the legend, "w/" and "w/o MM" refer to the calculations with or without the MM subsystem, respectively. Transition dipole moments computed for the QM clusters, QM subsystems, and QM/MM models are compared in Section S3 of the Supporting Information.

DOI: 10.1021/acs.jctc.5b00197
J. Chem. Theory Comput. 2015, 11, 3878−3894

Table 3. Dipole Moment (Debye) in the Ground and Excited Electronic States in the QM/MM Models Computed with the CASSCF Wave Function[a]

	S_0	S_1	S_2	ET
QM subsystem	QM/MM (QM)	QM/MM (QM)	QM/MM (QM)	QM/MM (QM)
LfQY-N	13.65 (10.26)	17.25 (11.00)	13.64 (11.72)	39.58 (37.59)
LfQY	10.46 (7.84)	15.48 (10.57)	10.99 (10.10)	37.66 (36.53)
LfQY-S	10.96 (7.42)	16.28 (10.90)	11.75 (19.80[b])	37.78 (20.58[b])
LfQY-M	9.39 (7.00)	14.28 (9.85)	9.99 (9.11)	35.52 (29.94)
LfQY-MS	9.88 (6.79)	15.09 (10.40)	10.77 (9.03)	35.64 (34.40)

[a]QM/MM dipole moments are compared to those of the isolated QM subsystem indicated in brackets. [b]Mixing of the S_2 and ET states at the state crossing

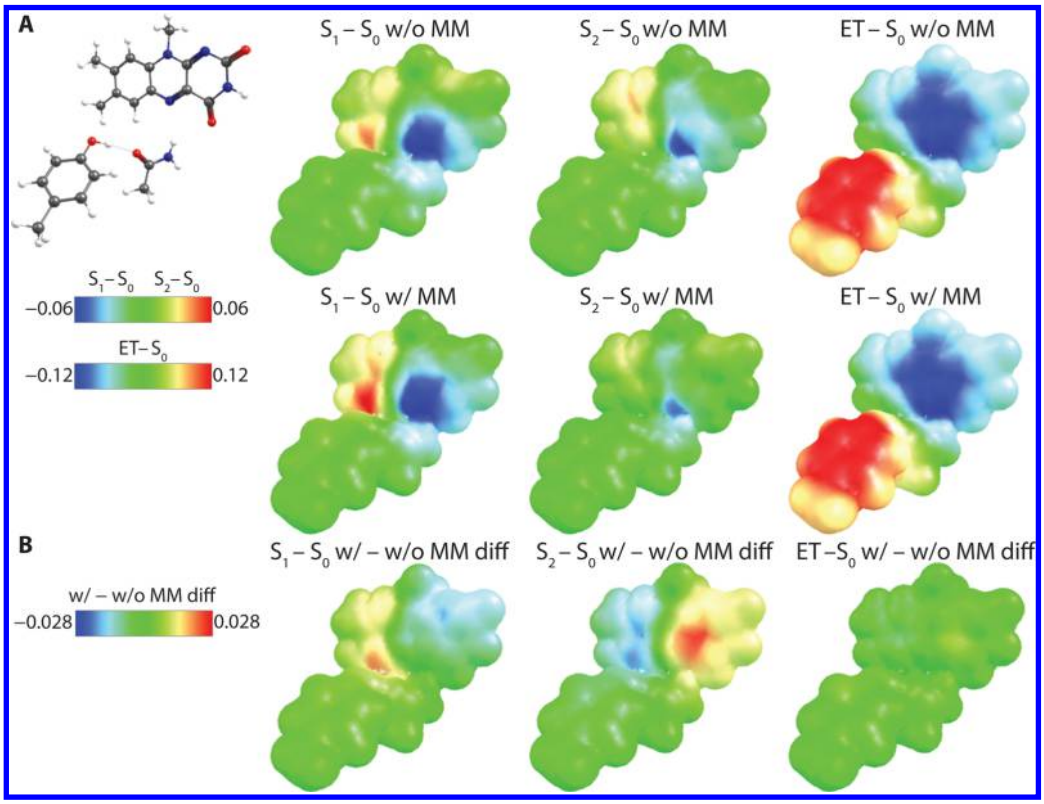


Figure 6. Charge transfer in the QM subsystem. (A) Difference (excited state minus S_0) molecular electrostatic potentials (MEPs) (au) of the LfQY QM subsystem in comparison to the LfQY/MM model. (B) Double difference MEPs demonstrate the difference polarization of the two states upon embedding. In (A) and (B), "w/" and "w/o MM" indicate the calculations with and without the MM subsystem, respectively.

subsystem contributes to the excitation energy shifts in two ways: through the difference in the wave function polarization energy of the two states as defined below and via the modification of the charge transfer magnitude, which in turn influences the interactions of the LfQY triad with the environment.

In the QM/MM model with electrostatic potential energy fitted (ESPF) coupling, the electrostatic interaction energy between the QM and MM subsystem is approximated using a classical expression eq 3.[50,68]

$$E_{\mathrm{QM/MM}}^{\mathrm{ESPF}} = \sum_{i,k} \frac{q_i Q_k}{r_{ik}} \tag{3}$$

where Q_k are the MM subsystem charges, and q_i are elements of the ESPF operator vector expectation values (i.e., the ESPF atomic charges) defined to fit the electrostatic integrals describing interactions between the QM and MM subsystems.[50,68] As the fitting is performed in the presence of the external MM charges, the ESPF charges incorporate the electron density polarization induced by the MM charges. Electrostatic embedding is introduced by adding the ESPF term (eq 3) to the Fock operator. At convergence, the total ESPF QM/MM energy incorporates the electrostatic interaction energy between the polarized QM subsystem ESPF charges and the MM subsystem charges and an electronic energy penalty for the polarization of the QM wave function by MM charges, which we call wave function polarization energy. From our QM/MM calculations, we directly obtain the electrostatic ESPF energy component (from now on called "ESPF energy") defined by eq 3. The wave function polarization energy component is instead computed as the difference between the total QM/MM energy and QM energy (respectively computed in presence and absence of electrostatic embedding) minus the ESPF energy. Our working definition of the wave function polarization energy is especially convenient in the case of using the ESPF operator QM/MM method, but it is different from

DOI: 10.1021/acs.jctc.5b00197
J. Chem. Theory Comput. 2015, 11, 3878−3894

the conventional definition.[90−92] Specifically, Gao and Xia[90] considered as solute−solvent polarization energy the electron density distortion energy (equivalent to our wave function polarization energy) plus a gain in the interaction energy between the polarized QM subsystem and MM subsystem over that of an "unpolarized" QM subsystem. In the ESPF operator QM/MM coupling scheme, the energy gain is computed through the electrostatic component, and it is included in the ESPF energy. Using our QM/MM energy decomposition into ESPF energy and wave function polarization energy, we obtain, for each given transition (e.g., S_0−ET), the components of the embedding energy shift $\Delta\varepsilon_{emb}$ (eq 4): the two-state ESPF energy difference $\Delta\varepsilon_{ESPF}$ (eq 5) and the two-state wave function polarization energy difference $\Delta\varepsilon_{wfpol}$ (eq 6).

$$\Delta\varepsilon_{emb} = \varepsilon_{QM/MM} - \varepsilon_{QM} \tag{4}$$

$$\Delta\varepsilon_{ESPF} = \sum_{i,k}\left(\frac{q_i^{ET}Q_k}{r_{ik}} - \frac{q_i^{S_0}Q_k}{r_{ik}}\right) = \sum_{i,k}\frac{\Delta q_i Q_k}{r_{ik}} \tag{5}$$

$$\Delta\varepsilon_{wfpol} = \varepsilon_{QM/MM} - \varepsilon_{QM} - \Delta\varepsilon_{ESPF} = \Delta\varepsilon_{emb} - \Delta\varepsilon_{ESPF} \tag{6}$$

where the excitation energy $\varepsilon_{QM/MM}$ is computed for the QM/MM system, and ε_{QM} is computed for the corresponding isolated QM subsystem. The two-state ESPF energy difference is proportional to the amount of charge transfer Δq_i of the excitation in eq 5. Notice that eq 5 compares with eq 1, which gives the Coulomb energy shift $\Delta\varepsilon_C$ for a QM cluster.

In the BLUF QM/MM models, the ESPF energy is negative and, therefore, stabilizes the system. In contrast, the wave function polarization energy is positive and is approximately one order of magnitude smaller than the ESPF energy. A more negative ESPF energy of S_0 than that of the excited state increases the excitation energy, while a higher wave function polarization energy of S_0 with respect to that of the excited state decreases the excitation energy. As our energy decomposition revealed, the embedding shift (eq 4), which is also referred to as the energy difference between the protein model and the "gas phase" model,[93] is defined not only by electrostatic interactions but also by wave function polarization. In particular, the embedding shifts of the spectroscopically observed S_0−S_1 and S_0−S_2 energies of flavin in our QM/MM models are dominated by the wave function polarization energy difference.

We found the following electrostatic ESPF energy shifts: a red shift for S_0−S_1 energy, no shift for S_0−S_2 energy, and a substantial blue shift for S_0−ET energy (Figure 7). Although the ESPF charges derived from the CASSCF and CASPT2 calculations differ (data not shown), the respective ESPF energy shifts are of similar magnitudes. The largest shift, as expected from the amount of charge transfer, is found for the S_0−ET transition. We note that for such a significant effect, the accuracy of both the ESPF charges and the MM charges may significantly contribute to the accuracy of the S_0−ET energy estimate. In Section S5 of the Supporting Information, we compare the Coulomb energy shifts defined by eq 1 computed for isolated QM subsystems with different charge models (Mulliken, ESPF, and RESP). We noticed that using the ESPF charges instead of the Mulliken charges increases the blue and decreases the red shift computed for the QM subsystem. The deviation should increase upon increasing the number of MM charges, i.e., if charges of the entire protein are included. At present, we cannot exclude that a charge model artifact affects the S_0−ET energy estimate in our QM/MM model. Scaling atomic charges to reduce the Coulomb energies when evaluating the protein electrostatic effect is, in fact, a widely used practice.[94,95]

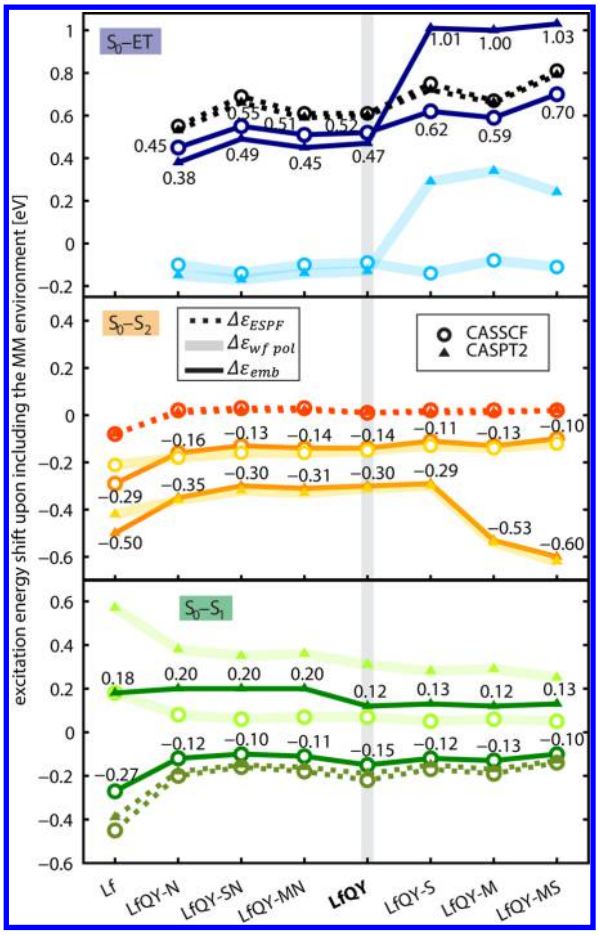


Figure 7. Energy decomposition analysis for the excitation energy shifts in the QM/MM model. The shifts indicated in the legend are explained in Subsection 3.6. The numbers report the $\Delta\varepsilon_{emb}$ shift at the CASSCF and CASPT2 levels. Note that the CASPT2 energy artifact at the S_2/ET state crossing in the QM subsystems LfQY-S, LfQY-M, and LfQY-MS increases the S_0−S_2 and S_0−ET wave function polarization energy shifts.

The wave function polarization shifts at the CASPT2 level are as follows: a significant blue shift for S_0−S_1, a significant red shift for S_0−S_2, and a small red shift for the S_0−ET energy (Figure 7). The shift increases when computed at the CASPT2 level causing the difference in the CASPT2 and CASSCF embedding shifts. Specifically, the blue shift of S_0−S_1 upon embedding computed with CASPT2 and MS-CASPT2 is caused by the increased polarization of the S_1 state. Unfortunately, the spurious behavior of the CASPT2 energies at the S_2/ET state crossing artificially increases the wave function polarization shift, complicating our analysis of the polarization effects upon enlarging the QM subsystem. However, if no side chains are included in the QM subsystem, the S_0−S_1 and S_0−S_2 wave function polarization shifts substantially increase in the absolute magnitude. We might expect a decrease in the two-state wave function polarization energy difference down to a certain value upon sufficient enlargement of the QM subsystem.

DOI: 10.1021/acs.jctc.5b00197
J. Chem. Theory Comput. 2015, 11, 3878−3894

We note that the magnitude of the S_0-S_2 energy in the QM/MM models is solely determined by the wave function polarization shift as the respective electrostatic ESPF component is eliminated by the larger S_0 polarization in comparison to S_2 (i.e., evaluated according to eq 6 and confirmed by changes of dipole moment in Table 3). In this case, the absence of intermolecular charge transfer of the electronic transition, causing the only negligible ESPF energy shift, is inconsistent with the results of the QM cluster showing a larger amount of S_0-S_2 charge transfer. In the absence of the charge transfer character, the QM/MM S_0-S_2 energy shows small variations by apoprotein modifications tested here (see also data in Sections S6 and S7 of the Supporting Information).

To conclude, the differences between the QM cluster and QM/MM excitation energies are determined by a number of factors. The decreased S_0-S_1 and S_0-S_2 energies obtained with the QM/MM model in comparison with the QM cluster are due to the different PT2 methods used (CASPT2 in comparison with XMCQDPT2). An additional decrease in the S_0-S_2 energy arises from the wave function polarization energy, i.e., the modest S_2 wave function polarization in comparison to the large S_0 wave function polarization in the QM/MM model. We note that because of these polarization effects, the charge transfer character of the S_0-S_2 transition is drastically diminished. Therefore, the QM/MM model does not reveal any protein effects on the S_0-S_2 energies governed by electrostatic interactions. In contrast, the S_0-ET energy increases in the QM/MM model as compared to the QM cluster because of the protein electrostatics. The blue electrostatic shift of the S_0-ET energy in the protein is to some extent compensated by red shifts due to wave function polarization and due to the shortened electron donor−acceptor distance in the QM/MM model.

3.7. Color and Redox Tuning in the QM/MM Model.

Protein electrostatic effects are essential for color tuning of rhodopsin and green fluorescent proteins because of the charge transfer character of their excited electronic states.[79,80,93,96] In particular, computational models[93,96] demonstrated that color changes of rhodopsin photoreceptors depend on stabilization/destabilization of the positive charge that in the ground state is localized on the protonated Schiff base region of the retinal chromophore and in the excited state delocalized toward retinal's β-ionone ring. In contrast to rhodopsin, flavin-containing photoreceptors do not integrate relevant charge localization on the chromophore region in the ground state. Consequently, the role of electrostatic effects is defined by stabilization/destabilization of more polar S_1, S_2, or even charge separated ET excited states. To document the effect of the protein electrostatics in the BLUF photoreceptor that features excited states with different amount of charge transfer, we prepared the LFQY/MM$^{\pm}$ series of models, in which we introduced an additional +1 and −1 MM charge on selected side chains in the vicinity of the flavin and tyrosine Y11. Accordingly, we selected the Y66 and H75 side chains near Y11 and the S30 side chain and the FMN phosphate group near flavin (Figure 8A). Upon the selected charge modification, significant variations were found for the S_0-ET energy, in contrast to small S_0-S_1 and S_0-S_2 energy changes (Figure 8B). Thus, the calculations show that the protein electrostatics substantially influences the redox potential but has more limited effect on the flavin color. The observed excitation energy shifts of the S_0-S_1 and S_0-ET transitions follow the trends of the ESPF energy change as demonstrated in Section S7 of the Supporting Information. The corresponding energy difference due to wave function polarization stays approximately constant for the selected charge modifications, although it varies more for the S_0-S_1 energy compared to the S_0-ET energy. At the same time the changes of the S_0-S_2 energy are defined by the wave function polarization energy difference of the two states.

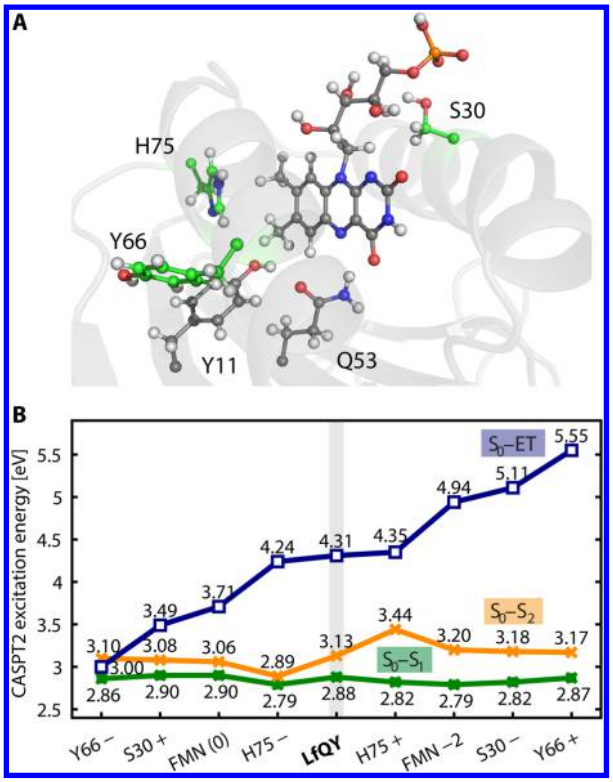


Figure 8. Excitation energies in the LfQY/MM model with the altered MM charges. (A) Cartoon representation of the flavin-binding pocket in PixD BLUF (prepared with Pymol[97]) with the green-colored residues for which the charge was changed by +1 or −1 au. (B) CASPT2 excitation energies of the LfQY/MM$^{\pm}$ models with the positively and negatively charged residues Y66, S30, H75, and neutral FMN(0) or negatively charged FMN−2 phosphate compared to the energies of the LfQY/MM model. All models have been computed with modified charge redistribution at the QM/MM boundary (see Computational Details).

The role of electrostatic interactions in determining small S_0-S_1 energy shifts in various mutated BLUF proteins were recently explored by Collette et al.,[94] who established a general rule by considering the difference excited-minus-ground state electrostatic potential of the S_0-S_1 transition. The mutation resulting in removal of an atom with a positive partial charge from the positive region of the difference electrostatic potential (i.e., close to flavin N1, C2, and O2 atoms; for atom labeling, see Scheme 1) results in the red-shifted absorbance maximum of the mutant, whereas removal of a negative atom from the same region results in the blue-shifted absorbance maximum. The inverse effect was predicated for the negative region of the difference electrostatic potential (i.e., close to flavin C7, N5, C4a, C4, and O4 atoms).[94] Our results are overall consistent with the rule that was described above. We find the S_0-S_1 energy decreases when the positive charge is removed from S30 located in the positive difference potential region close to N1 and C2 and the S_0-S_1 energy increases when removing the

DOI: 10.1021/acs.jctc.5b00197
J. Chem. Theory Comput. 2015, 11, 3878−3894

positive charge from H75 located in the negative difference potential region close to C7.

The large changes of the S_0−ET energy are governed by electrostatic stabilization/destabilization of the positive and negative charges on tyrosine Y11 and flavin, respectively. A negative charge close to flavin isoalloxazine group, for instance, a deprotonated FMN phosphate, destabilizes the flavin radical-anion and increases the S_0−ET energy (Figure 8B). It is interesting to note that the blue ESPF energy shift of the S_0−ET energy vanishes when the phosphate charge is neutralized in the FMN(0) model (Figure S6B, Supporting Information). Such charge neutralization can occur, for instance, by counterion binding or by substituting the N34 side chain of PixD BLUF with a positively charged arginine or histidine found in BlrB[39] or AppA[98] BLUF domains, respectively. Indeed, the BLUF domains from various organisms and their mutants show variations in the fluorescence lifetime, efficiency of the signaling state formation, and signaling state lifetime, which are linked to redox-potential modulations.[99] Yet, mutants of PixD BLUF with N34 substituted to histidine or arginine demonstrated longer excited flavin lifetime[99] (consistent with higher S_0−ET energy than in the wild type), suggesting that other factors than electrostatic interactions of these side chains with flavin radical-anion play a role in the observed phenomena.

At present, our results are qualitative and establish that the charge transfer characters of electronic transitions govern the excitation energy changes induced by different environments. Our calculations predict that BLUF domains exhibiting very similar flavin UV−vis absorbance spectra may vary significantly in flavin fluorescence lifetimes, controlled by the S_1−ET transition. Yet, to enable more quantitative predictions, instead of a single protein−solvent configuration, protein dynamics must be taken into account as was previously done, for instance, for bovine rhodopsin and its mutants.[93] While combining MD and QM/MM studies of BLUF proteins should be a subject of future work, MD studies of several BLUF domains revealed unstable in time interactions of the side chains with flavin[100,101] leading to significant deviations from the respective crystal structures. Rieff et al.[100] attributed this result to the neglect of polarization effects in the currently employed biomolecular force fields. Nonetheless, experimental studies[39,98,102,103] also demonstrated that some BLUF-domain constructs have a rather disordered flavin-binding site in solution, supporting the hypothesis that dynamic flavin−apoprotein interactions constitute an intrinsic feature (see also discussion in ref 23). Thus, accurate excitation energy computations must be combined with careful structural analysis and validation of MD protocols in studies aiming at comparing homologous BLUF proteins. In addition to the role of electrostatic effects, our current results indicate that variations in the distribution of the donor−acceptor distance may significantly influence the S_0−ET and S_1−ET excitation energies. Therefore, applying geometry constraints/restraints at the flavin-binding site in MD models may yield biased results.

4. CONCLUSIONS

Above, we have presented an extensive computational study of the vertical excitation energies in the photosensory protein BLUF. Our focus was on establishing how the energies of excited flavin and of electron transfer from tyrosine to flavin depend on interactions with the apoprotein. Accordingly, we compared the performance of two alternative approaches: a QM cluster and QM/MM model. Our computations revealed that the trends in the BLUF excitation energies, independent of whether they are computed with the QM cluster or QM/MM model, are governed by a single fundamental property: the amount of charge transfer in the electronic transition. Our goal was not only to find out which properties are confirmed by both models, but also to describe the differences between these common alternative approaches. The comparison between the results obtained with our specific QM cluster and QM/MM models provides information on factors such as the geometry of the active site, performance of the quantum mechanical methods, and the role of electrostatic, polarization and charge transfer interactions.

Upon geometry optimization, the flavin−tyrosine distance of the QM cluster becomes longer as compared to the crystal structure. In contrast, the QM/MM model yields a more compact geometry with distances close to those in the crystal structure. This structural difference does not influence the S_0−S_1 and S_0−S_2 energies but impacts the S_0−ET energy, which decreases by ca. 0.3 eV upon decreasing the flavin−tyrosine distance in the QM/MM optimized geometry. Therefore, a QM/MM geometry, which is close to the respective X-ray crystal structure, may yield facilitated flavin fluorescence quenching and photoreduction as compared to a cluster model geometry.

We found substantial systematic discrepancies in the computed XMCQDPT2 and MS-CASPT2 excitation energies. There is a pronounced decrease in the S_0−S_1 and S_0−S_2 MS-CASPT2 values with respect to the corresponding XMCQDPT2 values (0.2−0.3 eV), while the difference in the S_0−ET estimate of the two methods is significantly smaller (<0.05 eV). However, upon increasing the size of the QM subsystem and stabilization of the ET state, the discrepancy in the S_0−ET estimate of the two methods becomes rather large (0.5 eV). We assign this discrepancy to incorrect MS-CASPT2 energy values, when the crossing electronic states mix in the zeroth-order CASSCF wave function (see also ref 89). Interestingly, coupling between the QM and MM subsystems in the QM/MM model lifts the degeneracy and removes the artifact.

Both the QM cluster and QM/MM model demonstrate that the impact of the molecular environment on the excitation energies depends on the amount of charge transfer (i.e., the change of the dipole moment) of the electronic transition. In BLUF, as well as in other oxidized flavoproteins, the S_0−ET energy and, in turn, the protein redox properties are most sensitive to the interaction with the apoprotein, whereas the S_0−S_1 energy defining the flavin color displays only small variation. The more pronounced is the excitation charge transfer character, the larger is the role of electrostatic interactions but also of charge transfer with the apoprotein. More specifically, the electrostatic interactions with the MM environment increase the S_0−ET energy by 0.6 eV but decrease the S_0−S_1 energy only by ca. 0.2 eV. In the QM cluster model, which includes only a small part of the environment compared to the QM/MM model, the electrostatic interactions decreases the S_0−ET energy by 0.1 eV and transfer of charge between the triad and its environment (absent in the QM/MM model if the QM subsystem includes only the triad) further decreases the energy by ca. 0.1 eV. Overall, we observed that the interactions with the environment in the QM/MM model inhibit flavin photoreduction (despite the opposite effect of the optimized geometry mentioned above), while the interactions in the QM

DOI: 10.1021/acs.jctc.5b00197
J. Chem. Theory Comput. 2015, 11, 3878−3894

cluster facilitate the photoreduction. Eventually, our two current models would yield a significantly different flavin fluorescence lifetime due to the S_1−ET transition (shorter lifetime according to the QM cluster S_1−ET energy and longer lifetime according to the QM/MM model S_1−ET energy).

The wave function polarization in the QM/MM model (observed, for instance, as the dipole moment change upon electrostatic embedding) influences the magnitude of charge transfer in the electronic transition and consequently modifies the excitation energy. The wave function polarization energy difference of the two states defined in Subsection 3.6 is smallest for the most polar S_0−ET transition (decreases its energy by 0.15 eV), while it is substantial for the transitions within the flavin chromophore (increases the S_0−S_1energy by 0.3 eV and decreases the S_0−S_2 energy by 0.3 eV). Moreover, the polarization increases and decreases the amount of charge transfer in the S_0−S_1 and S_0−S_2 transition, respectively, modifying the electrostatic interaction. The wave function polarization might introduce artifacts in several ways, and the most striking artifact we observed in our study is the absence of the charge transfer character in the S_0−S_2 excitation in the QM/MM model in disagreement with the results of our QM cluster models.

Finally, we note that in addition to the effects analyzed in our study, the results of QM/MM computations depend on the preparation of the protein structure and the protein model parameters (such as, ionization status of the residues, geometrical constraints, force field, definition of the QM/MM boundary and others). Evaluating the impact from these factors requires a significant effort; hence, in practice, keeping all these parameters constant eliminates their contributions to the final relative energies. Our current QM/MM model is mostly suitable for evaluating local interactions, e.g., effects of single side chains, under the assumption that the protein crystal structure provides a reasonable geometry for the local interactions in question. For this particular application, the local effects are accurately estimated in the QM cluster. To ensure higher accuracy as well as consistency with QM/MM models and protein structures, a large part of the active-site environment should be included in the cluster model and rather conservative geometrical constraints should be used upon geometry optimization.

■ ASSOCIATED CONTENT

Supporting Information

Section S1, List of models; Section S2, Estimation of BSSE in the XMCQDPT2 calculations; Section S3, Transition dipole moments; Section S4, Electrostatic potential maps of the BLUF flavin binding site; Section S5, Estimation of Coulomb and nonelectrostatic interactions in the QM subsystems; Section S6, Energy decomposition of the LfQY/MM^{off} models; and Section S7, Energy decomposition of the LfQY/$MM^{+1/-1}$ models. The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jctc.5b00197.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail:Tatjana.Domratcheva@mpimf-heidelberg.mpg.de.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We are very grateful to Chris Roome (MPImF Heidelberg) for excellent IT support; to Alex Granovsky (Firefly project) for computational advice, valuable discussion, and critical reading of the manuscript; to Nicolas Ferré (Université de Provence) for help with MOLCAS/Tinker calculations; to Federico Melaccio (University of Siena) and Samer Gozem, Igor Schapiro, and Mark Huntress (Bowling Green State University) for helpful discussions and advice with MOLCAS/Tinker calculations. T.D. and A.U. are grateful to Ilme Schlichting (MPImF Heidelberg) for kind support and acknowledge financial support from the Minerva Program of the Max Planck Society to T.D. and from the Boehringer Ingelheim Fonds and the Heidelberg Graduate School of Mathematical and Computational Methods for the Sciences (HGSMathComp) to A.U. M.O. acknowledges partial support from the National 1051 Science Foundation (Grant no. CHE-1152070), Human 1052 Frontier Science Program Organization (Grant RGP0049/1053 2012CHE09-56776), EU-FP7 (Marie-Curie PIOF-GA-1054 2012-332233) and from the University of Strasbourg Institute for Advanced Study (USIAS).

■ REFERENCES

(1) Losi, A.; Gärtner, W. Old Chromophores, New Photoactivation Paradigms, Trendy Applications: Flavins in Blue Light-Sensing Photoreceptors. *Photochem. Photobiol.* **2011**, *87*, 491−510.
(2) Losi, A.; Gärtner, W. The Evolution of Flavin-Binding Photoreceptors: An Ancient Chromophore Serving Trendy Blue-Light Sensors. *Annu. Rev. Plant Biol.* **2012**, *63*, 49−72.
(3) Cashmore, A. R.; Jarillo, J. A.; Wu, Y.-J.; Liu, D. Cryptochromes: Blue Light Receptors for Plants and Animals. *Science* **1999**, *284*, 760−765.
(4) Christie, J. M.; Gawthorne, J.; Young, G.; Fraser, N. J.; Roe, A. J. LOV to BLUF: Flavoprotein Contributions to the Optogenetic Toolkit. *Mol. Plant* **2012**, *5*, 533−544.
(5) Herrou, J.; Crosson, S. Function, Structure and Mechanism of Bacterial Photosensory LOV Proteins. *Nat. Rev. Microbiol.* **2011**, *9*, 713−723.
(6) Wingen, M.; Potzkei, J.; Endres, S.; Casini, G.; Rupprecht, C.; Fahlke, C.; Krauss, U.; Jaeger, K.-E.; Drepper, T.; Gensch, T. The Photophysics of LOV-Based Fluorescent Proteins − New Tools for Cell Biology. *Photochem. Photobiol. Sci.* **2014**, *13*, 875−883.
(7) Gauden, M.; Yeremenko, S.; Laan, W.; van Stokkum, I. H. M.; Ihalainen, J. A.; van Grondelle, R.; Hellingwerf, K. J.; Kennis, J. T. M. Photocycle of the Flavin-Binding Photoreceptor AppA, a Bacterial Transcriptional Antirepressor of Photosynthesis Genes. *Biochemistry* **2005**, *44*, 3653−3662.
(8) Gauden, M.; van Stokkum, I. H. M.; Key, J. M.; Lührs, D. C.; van Grondelle, R.; Hegemann, P.; Kennis, J. T. M. Hydrogen-Bond Switching through a Radical Pair Mechanism in a Flavin-Binding Photoreceptor. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 10895−10900.
(9) Zirak, P.; Penzkofer, A.; Schiereis, T.; Hegemann, P.; Jung, A.; Schlichting, I. Absorption and Fluorescence Spectroscopic Characterization of BLUF Domain of AppA from Rhodobacter Sphaeroides. *Chem. Phys.* **2005**, *315*, 142−154.
(10) Fujiyoshi, S.; Hirano, M.; Matsushita, M.; Iseki, M.; Watanabe, M. Structural Change of a Cofactor Binding Site of Flavoprotein Detected by Single-Protein Fluorescence Spectroscopy at 1.5 K. *Phys. Rev. Lett.* **2011**, *106*, 078101.
(11) Sun, M.; Moore, T. A.; Song, P.-S. Molecular Luminescence Studies of Flavines. I. Excited States of Flavines. *J. Am. Chem. Soc.* **1972**, *94*, 1730−1740.
(12) Massey, V. The Chemical and Biological Versatility of Riboflavin. *Biochem. Soc. Trans.* **2000**, *28*, 283−296.
(13) Domratcheva, T.; Fedorov, R.; Schlichting, I. Analysis of the Primary Photocycle Reactions Occurring in the Light, Oxygen, and

DOI: 10.1021/acs.jctc.5b00197
J. Chem. Theory Comput. 2015, 11, 3878−3894

Voltage Blue-Light Receptor by Multiconfigurational Quantum-Chemical Methods. *J. Chem. Theory Comput.* **2006**, *2*, 1565−1574.
(14) Udvarhelyi, A.; Domratcheva, T. Photoreaction in BLUF Receptors: Proton-Coupled Electron Transfer in the Flavin-Gln-Tyr System. *Photochem. Photobiol.* **2011**, *87*, 554−563.
(15) Sadeghian, K.; Schütz, M. On the Photophysics of Artificial Blue-Light Photoreceptors: An Ab Initio Study on a Flavin-Based Dye Dyad at the Level of Coupled-Cluster Response Theory. *J. Am. Chem. Soc.* **2007**, *129*, 4068−4074.
(16) Sadeghian, K.; Bocola, M.; Schütz, M. A Conclusive Mechanism of the Photoinduced Reaction Cascade in Blue Light Using Flavin Photoreceptors. *J. Am. Chem. Soc.* **2008**, *130*, 12501−12513.
(17) Sadeghian, K.; Bocola, M.; Schütz, M. A QM/MM Study on the Fast Photocycle of Blue Light Using Flavin Photoreceptors in Their Light-Adapted/active Form. *Phys. Chem. Chem. Phys.* **2010**, *12*, 8840−8846.
(18) Sadeghian, K.; Bocola, M.; Merz, T.; Schütz, M. Theoretical Study on the Repair Mechanism of the (6−4) Photolesion by the (6−4) Photolyase. *J. Am. Chem. Soc.* **2010**, *132*, 16285−16295.
(19) Solov'yov, I. A.; Domratcheva, T.; Moughal Shahi, A. R.; Schulten, K. Decrypting Cryptochrome: Revealing the Molecular Identity of the Photoactivation Reaction. *J. Am. Chem. Soc.* **2012**, *134*, 18046−18052.
(20) Solov'yov, I. A.; Domratcheva, T.; Schulten, K. Separation of Photo-Induced Radical Pair in Cryptochrome to a Functionally Critical Distance. *Sci. Rep.* **2014**, *4*, 3845.
(21) Domratcheva, T. Neutral Histidine and Photoinduced Electron Transfer in DNA Photolyases. *J. Am. Chem. Soc.* **2011**, *133*, 18172−18182.
(22) Silva-Junior, M. R.; Mansurova, M.; Gärtner, W.; Thiel, W. Photophysics of Structurally Modified Flavin Derivatives in the Blue-Light Photoreceptor YtvA: A Combined Experimental and Theoretical Study. *ChemBioChem* **2013**, *14*, 1648−1661.
(23) Domratcheva, T.; Udvarhelyi, A.; Shahi, A. R. M. Computational Spectroscopy, Dynamics, and Photochemistry of Photosensory Flavoproteins. In *Flavins and Flavoproteins: Methods in Molecular Biology*; Weber, S., Schleicher, E., Eds.; Springer: New York, 2014; Chapter 10, pp 191−228.
(24) Neiss, C.; Saalfrank, P.; Parac, M.; Grimme, S. Quantum Chemical Calculation of Excited States of Flavin-Related Molecules. *J. Phys. Chem. A* **2003**, *107*, 140−147.
(25) Salzmann, S.; Tatchen, J.; Marian, C. M. The Photophysics of Flavins: What Makes the Difference between Gas Phase and Aqueous Solution? *J. Photochem. Photobiol., A* **2008**, *198*, 221−231.
(26) Khrenova, M. G.; Nemukhin, A. V.; Grigorenko, B. L.; Krylov, A. I.; Domratcheva, T. M. Quantum Chemistry Calculations Provide Support to the Mechanism of the Light-Induced Structural Changes in the Flavin-Binding Photoreceptor Proteins. *J. Chem. Theory Comput.* **2010**, *6*, 2293−2302.
(27) Udvarhelyi, A.; Domratcheva, T. Glutamine Rotamers in BLUF Photoreceptors: A Mechanistic Reappraisal. *J. Phys. Chem. B* **2013**, *117*, 2888−2897.
(28) Peach, M. J. G.; Benfield, P.; Helgaker, T.; Tozer, D. J. Excitation Energies in Density Functional Theory: An Evaluation and a Diagnostic Test. *J. Chem. Phys.* **2008**, *128*, 044118.
(29) Yanai, T.; Tew, D. P.; Handy, N. C. A New Hybrid Exchange−correlation Functional Using the Coulomb-Attenuating Method (CAM-B3LYP). *Chem. Phys. Lett.* **2004**, *393*, 51−57.
(30) Peach, M. J. G.; Tozer, D. J. Overcoming Low Orbital Overlap and Triplet Instability Problems in TDDFT. *J. Phys. Chem. A* **2012**, *116*, 9783−9789.
(31) Siegbahn, P. E. M.; Himo, F. The Quantum Chemical Cluster Approach for Modeling Enzyme Reactions. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2011**, *1*, 323−336.
(32) Kamerlin, S. C. L.; Haranczyk, M.; Warshel, A. Progress in Ab Initio QM/MM Free-Energy Simulations of Electrostatic Energies in Proteins: Accelerated QM/MM Studies of pKa, Redox Reactions and Solvation Free Energies. *J. Phys. Chem. B* **2009**, *113*, 1253−1272.
(33) Lin, H.; Truhlar, D. G. QM/MM: What Have We Learned, Where Are We, and Where Do We Go from Here? *Theor. Chem. Acc.* **2007**, *117*, 185−199.
(34) Senn, H. M.; Thiel, W. QM/MM Methods for Biomolecular Systems. *Angew. Chem., Int. Ed.* **2009**, *48*, 1198−1229.
(35) Ananikov, V. P.; Musaev, D. G.; Morokuma, K. Real Size of Ligands, Reactants and Catalysts: Studies of Structure, Reactivity and Selectivity by ONIOM and Other Hybrid Computational Approaches. *J. Mol. Catal. A: Chem.* **2010**, *324*, 104−119.
(36) Monari, A.; Rivail, J.-L.; Assfeld, X. Theoretical Modeling of Large Molecular Systems. Advances in the Local Self Consistent Field Method for Mixed Quantum Mechanics/Molecular Mechanics Calculations. *Acc. Chem. Res.* **2013**, *46*, 596−603.
(37) Gao, J. Methods and Applications of Combined Quantum Mechanical and Molecular Mechanical Potentials. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; John Wiley & Sons, Inc., 1996; pp 119−185.
(38) Gomelsky, M.; Klug, G. BLUF: A Novel FAD-Binding Domain Involved in Sensory Transduction in Microorganisms. *Trends Biochem. Sci.* **2002**, *27*, 497−500.
(39) Jung, A.; Domratcheva, T.; Tarutina, M.; Wu, Q.; Ko, W.; Shoeman, R. L.; Gomelsky, M.; Gardner, K. H.; Schlichting, I. Structure of a Bacterial BLUF Photoreceptor: Insights into Blue Light-Mediated Signal Transduction. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 12350−12355.
(40) Kita, A.; Okajima, K.; Morimoto, Y.; Ikeuchi, M.; Miki, K. Structure of a Cyanobacterial BLUF Protein, Tll0078, Containing a Novel FAD-Binding Blue Light Sensor Domain. *J. Mol. Biol.* **2005**, *349*, 1−9.
(41) Yuan, H.; Anderson, S.; Masuda, S.; Dragnea, V.; Moffat, K.; Bauer, C. Crystal Structures of the Synechocystis Photoreceptor Slr1694 Reveal Distinct Structural States Related to Signaling. *Biochemistry* **2006**, *45*, 12687−12694.
(42) Barends, T. R. M.; Hartmann, E.; Griese, J. J.; Beitlich, T.; Kirienko, N. V.; Ryjenkov, D. A.; Reinstein, J.; Shoeman, R. L.; Gomelsky, M.; Schlichting, I. Structure and Mechanism of a Bacterial Light-Regulated Cyclic Nucleotide Phosphodiesterase. *Nature* **2009**, *459*, 1015−1018.
(43) Domratcheva, T.; Grigorenko, B. L.; Schlichting, I.; Nemukhin, A. V. Molecular Models Predict Light-Induced Glutamine Tautomerization in BLUF Photoreceptors. *Biophys. J.* **2008**, *94*, 3872−3879.
(44) Roos, B. O. The Complete Active Space Self-Consistent Field Method and Its Applications in Electronic Structure Calculations. In *Advances in Chemical Physics*; Lawley, K. P., Ed.; John Wiley & Sons, Inc., 1987; pp 399−445.
(45) Granovsky, A. A. Extended Multi-Configuration Quasi-Degenerate Perturbation Theory: The New Approach to Multi-State Multi-Reference Perturbation Theory. *J. Chem. Phys.* **2011**, *134*, 214113−214114.
(46) Okajima, K.; Yoshihara, S.; Fukushima, Y.; Geng, X.; Katayama, M.; Higashi, S.; Watanabe, M.; Sato, S.; Tabata, S.; Shibata, Y.; Itoh, S.; Ikeuchi, M. Biochemical and Functional Characterization of BLUF-Type Flavin-Binding Proteins of Two Species of Cyanobacteria. *J. Biochem.* **2005**, *137*, 741−750.
(47) Finley, J.; Malmqvist, P.-Å.; Roos, B. O.; Serrano-Andrés, L. The Multi-State CASPT2Method. *Chem. Phys. Lett.* **1998**, *288*, 299−306.
(48) Granovsky, A. A. Firefly Version 8.0. http://classic.chem.msu.su/gran/firefly/index.html (accessed June 2015).
(49) Aquilante, F.; De Vico, L.; Ferré, N.; Ghigo, G.; Malmqvist, P.-åke; Neogrády, P.; Pedersen, T. B.; Pitoňák, M.; Reiher, M.; Roos, B. O.; Serrano-Andrés, L.; Urban, M.; Veryazov, V.; Lindh, R. MOLCAS 7: The Next Generation. *J. Comput. Chem.* **2010**, *31*, 224−247.
(50) Ferre, N.; Angyan, J. G. Approximate Electrostatic Interaction Operator for QM/MM Calculations. *Chem. Phys. Lett.* **2002**, *356*, 331−339.
(51) Baker, J.; Kessi, A.; Delley, B. The Generation and Use of Delocalized Internal Coordinates in Geometry Optimization. *J. Chem. Phys.* **1996**, *105*, 192−212.

DOI: 10.1021/acs.jctc.5b00197
J. Chem. Theory Comput. 2015, 11, 3878−3894

(52) Witek, H. A.; Choe, Y.-K.; Finley, J. P.; Hirao, K. Intruder State Avoidance Multireference Møller−Plesset Perturbation Theory. *J. Comput. Chem.* **2002**, *23*, 957−965.
(53) Boys, S. F.; Bernardi, F. The Calculation of Small Molecular Interactions by the Differences of Separate Total Energies. Some Procedures with Reduced Errors. *Mol. Phys.* **1970**, *19*, 553−566.
(54) Chemcraft version 1.6. http://www.chemcraftprog.com (accessed June 2015).
(55) 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. General Atomic and Molecular Electronic Structure System. *J. Comput. Chem.* **1993**, *14*, 1347−1363.
(56) Feller, D. The Role of Databases in Support of Computational Chemistry Calculations. *J. Comput. Chem.* **1996**, *17*, 1571−1586.
(57) Schuchardt, K. L.; Didier, B. T.; Elsethagen, T.; Sun, L.; Gurumoorthi, V.; Chase, J.; Li, J.; Windus, T. L. Basis Set Exchange: A Community Database for Computational Sciences. *J. Chem. Inf. Model.* **2007**, *47*, 1045−1052.
(58) Pronk, S.; Páll, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; van der Spoel, D.; Hess, B.; Lindahl, E. GROMACS 4.5: A High-Throughput and Highly Parallel Open Source Molecular Simulation Toolkit. *Bioinformatics* **2013**, *29*, 845−854.
(59) Oostenbrink, C.; Villa, A.; Mark, A. E.; Van Gunsteren, W. F. A Biomolecular Force Field Based on the Free Enthalpy of Hydration and Solvation: The GROMOS Force-Field Parameter Sets 53A5 and 53A6. *J. Comput. Chem.* **2004**, *25*, 1656−1676.
(60) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. Interaction Models for Water in Relation to Protein Hydration. In *Intermolecular Forces*; Pullman, B., Ed.; D. Reidel Publishing Company: Dordrecht, 1981; pp 331−342.
(61) Ponder, J. W.; Richards, F. M. An Efficient Newton-like Method for Molecular Mechanics Energy Minimization of Large Molecules. *J. Comput. Chem.* **1987**, *8*, 1016−1024.
(62) Gozem, S.; Melaccio, F.; Luk, H. L.; Rinaldi, S.; Olivucci, M. Learning from Photobiology How to Design Molecular Devices Using a Computer. *Chem. Soc. Rev.* **2014**, *43*, 4019−4036.
(63) El-Khoury, P. Z.; Schapiro, I.; Huntress, M.; Melaccio, F.; Gozem, S.; Frutos, L. M.; Olivucci, M. Computational Photochemistry and Photobiology. In *CRC Handbook of Organic Photochemistry and Photobiology*; Griesbeck, A. G., Oelgemöller, M., Ghetti, F., Eds.; Taylor & Francis: Boca Raton, FL, 2012; pp 1029−1056.
(64) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.
(65) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926−935.
(66) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. Automatic Atom Type and Bond Type Perception in Molecular Mechanical Calculations. *J. Mol. Graphics Modell.* **2006**, *25*, 247−260.
(67) Singh, U. C.; Kollman, P. A. A Combined Ab Initio Quantum Mechanical and Molecular Mechanical Method for Carrying out Simulations on Complex Molecular Systems: Applications to the CH3Cl + Cl− Exchange Reaction and Gas Phase Protonation of Polyethers. *J. Comput. Chem.* **1986**, *7*, 718−730.
(68) Melaccio, F.; Olivucci, M.; Lindh, R.; Ferré, N. Unique QM/MM Potential Energy Surface Exploration Using Microiterations. *Int. J. Quantum Chem.* **2011**, *111*, 3339−3346.
(69) Forsberg, N.; Malmqvist, P.-Å. Multiconfiguration Perturbation Theory with Imaginary Level Shift. *Chem. Phys. Lett.* **1997**, *274*, 196−204.
(70) Ghigo, G.; Roos, B. O.; Malmqvist, P.-Å. A Modified Definition of the Zeroth-Order Hamiltonian in Multiconfigurational Perturbation Theory (CASPT2). *Chem. Phys. Lett.* **2004**, *396*, 142−149.
(71) Olsen, S.; Lamothe, K.; Martínez, T. J. Protonic Gating of Excited-State Twisting and Charge Localization in GFP Chromophores: A Mechanistic Hypothesis for Reversible Photoswitching. *J. Am. Chem. Soc.* **2010**, *132*, 1192−1193.
(72) Coe, J. D.; Martínez, T. J. Ab Initio Molecular Dynamics of Excited-State Intramolecular Proton Transfer around a Three-State Conical Intersection in Malonaldehyde. *J. Phys. Chem. A* **2006**, *110*, 618−630.
(73) Levine, B. G.; Martínez, T. J. Ab Initio Multiple Spawning Dynamics of Excited Butadiene: Role of Charge Transfer. *J. Phys. Chem. A* **2009**, *113*, 12815−12824.
(74) Veryazov, V.; Malmqvist, P. Å.; Roos, B. O. How to Select Active Space for Multiconfigurational Quantum Chemistry? *Int. J. Quantum Chem.* **2011**, *111*, 3329−3338.
(75) Khrenova, M. G.; Domratcheva, T.; Schlichting, I.; Grigorenko, B. L.; Nemukhin, A. V. Computational Characterization of Reaction Intermediates in the Photocycle of the Sensory Domain of the AppA Blue Light Photoreceptor. *Photochem. Photobiol.* **2011**, *87*, 564−573.
(76) Mo, Y.; Bao, P.; Gao, J. Energy Decomposition Analysis Based on a Block-Localized Wavefunction and Multistate Density Functional Theory. *Phys. Chem. Chem. Phys.* **2011**, *13*, 6760−6775.
(77) Gordon, M. S.; Smith, Q. A.; Xu, P.; Slipchenko, L. V. Accurate First Principles Model Potentials for Intermolecular Interactions. *Annu. Rev. Phys. Chem.* **2013**, *64*, 553−578.
(78) Zhao, Z.; Rogers, D. M.; Beck, T. L. Polarization and Charge Transfer in the Hydration of Chloride Ions. *J. Chem. Phys.* **2010**, *132*, 014502.
(79) Hasegawa, J.; Fujimoto, K. J.; Nakatsuji, H. Color Tuning in Photofunctional Proteins. *ChemPhysChem* **2011**, *12*, 3106−3115.
(80) Zhou, X.; Sundholm, D.; Wesołowski, T. A.; Kaila, V. R. I. Spectral Tuning of Rhodopsin and Visual Cone Pigments. *J. Am. Chem. Soc.* **2014**, *136*, 2723−2726.
(81) Hasegawa, J.; Yanai, K.; Ishimura, K. Quantum Mechanical Molecular Interactions for Calculating the Excitation Energy in Molecular Environments: A First-Order Interacting Space Approach. *ChemPhysChem* **2015**, *16*, 305−311.
(82) Neugebauer, J. Subsystem-Based Theoretical Spectroscopy of Biomolecules and Biomolecular Assemblies. *ChemPhysChem* **2009**, *10*, 3148−3173.
(83) Humbert-Droz, M.; Zhou, X.; Shedge, S. V.; Wesolowski, T. A. How to Choose the Frozen Density in Frozen-Density Embedding Theory-Based Numerical Simulations of Local Excitations? *Theor. Chem. Acc.* **2014**, *133*, 1−20.
(84) Dresselhaus, T.; Neugebauer, J.; Knecht, S.; Keller, S.; Ma, Y.; Reiher, M. Self-Consistent Embedding of Density-Matrix Renormalization Group Wavefunctions in a Density Functional Environment. *J. Chem. Phys.* **2015**, *142*, 044111.
(85) Lee, A. J.; Rick, S. W. The Effects of Charge Transfer on the Properties of Liquid Water. *J. Chem. Phys.* **2011**, *134*, 184507.
(86) Mayhall, N. J.; Raghavachari, K. Charge Transfer Across ONIOM QM:QM Boundaries: The Impact of Model System Preparation. *J. Chem. Theory Comput.* **2010**, *6*, 3131−3136.
(87) Jovan Jose, K. V.; Raghavachari, K. Electrostatic Potential-Based Method of Balancing Charge Transfer Across ONIOM QM:QM Boundaries. *J. Chem. Theory Comput.* **2014**, *10*, 4351−4359.
(88) Isborn, C. M.; Götz, A. W.; Clark, M. A.; Walker, R. C.; Martínez, T. J. Electronic Absorption Spectra from MM and Ab Initio QM/MM Molecular Dynamics: Environmental Effects on the Absorption Spectrum of Photoactive Yellow Protein. *J. Chem. Theory Comput.* **2012**, *8*, 5092−5106.
(89) Gozem, S.; Huntress, M.; Schapiro, I.; Lindh, R.; Granovsky, A. A.; Angeli, C.; Olivucci, M. Dynamic Electron Correlation Effects on the Ground State Potential Energy Surface of a Retinal Chromophore Model. *J. Chem. Theory Comput.* **2012**, *8*, 4069−4080.
(90) Gao, J.; Xia, X. A Priori Evaluation of Aqueous Polarization Effects Through Monte Carlo QM-MM Simulations. *Science* **1992**, *258*, 631−635.
(91) Thompson, M. A. Hybrid Quantum Mechanical/Molecular Mechanical Force Field Development for Large Flexible Molecules: A

DOI: 10.1021/acs.jctc.5b00197
J. Chem. Theory Comput. 2015, 11, 3878−3894

Molecular Dynamics Study of 18-Crown-6. *J. Phys. Chem.* **1995**, *99*, 4794−4804.
(92) Thompson, M. A.; Schenter, G. K. Excited States of the Bacteriochlorophyll *b* Dimer of *Rhodopseudomonas viridis*: A QM/MM study of the Photosynthetic Reaction Center That Includes MM Polarization. *J. Phys. Chem.* **1995**, *99*, 6374−6386.
(93) Rajamani, R.; Lin, Y.-L.; Gao, J. The Opsin Shift and Mechanism of Spectral Tuning in Rhodopsin. *J. Comput. Chem.* **2011**, *32*, 854−865.
(94) Collette, F.; Renger, T.; Schmidt am Busch, M. Revealing the Functional States in the Active Site of BLUF Photoreceptors from Electrochromic Shift Calculations. *J. Phys. Chem. B* **2014**, *118*, 11109−11119.
(95) Lüdemann, G.; Woiczikowski, P. B.; Kubař, T.; Elstner, M.; Steinbrecher, T. B. Charge Transfer in E. Coli DNA Photolyase: Understanding Polarization and Stabilization Effects via QM/MM Simulations. *J. Phys. Chem. B* **2013**, *117*, 10769−10778.
(96) Melaccio, F.; Ferré, N.; Olivucci, M. Quantum Chemical Modeling of Rhodopsin Mutants Displaying Switchable Colors. *Phys. Chem. Chem. Phys.* **2012**, *14*, 12485−12495.
(97) *The PyMOL Molecular Graphics System*, Version 1.3r1; Schrödinger, LLC, 2010.
(98) Jung, A.; Reinstein, J.; Domratcheva, T.; Shoeman, R. L.; Schlichting, I. Crystal Structures of the AppA BLUF Domain Photoreceptor Provide Insights into Blue Light-Mediated Signal Transduction. *J. Mol. Biol.* **2006**, *362*, 717−732.
(99) Mathes, T.; van Stokkum, I. H. M.; Stierl, M.; Kennis, J. T. M. Redox Modulation of Flavin and Tyrosine Determines Photoinduced Proton-Coupled Electron Transfer and Photoactivation of BLUF Photoreceptors. *J. Biol. Chem.* **2012**, *287*, 31725−31738.
(100) Rieff, B.; Bauer, S.; Mathias, G.; Tavan, P. DFT/MM Description of Flavin IR Spectra in BLUF Domains. *J. Phys. Chem. B* **2011**, *115*, 11239−11253.
(101) Meier, K.; Thiel, W.; van Gunsteren, W. F. On the Effect of a Variation of the Force Field, Spatial Boundary Condition and Size of the QM Region in QM/MM MD Simulations. *J. Comput. Chem.* **2012**, *33*, 363−378.
(102) Wu, Q.; Ko, W.-H.; Gardner, K. H. Structural Requirements for Key Residues and Auxiliary Portions of a BLUF Domain. *Biochemistry* **2008**, *47*, 10271−10280.
(103) Grinstead, J. S.; Hsu, S. T.; Laan, W.; Bonvin, A. M.; Hellingwerf, K. J.; Boelens, R.; Kapstein, R. The Solution Structure of the AppA BLUF Domain: Insight into the Mechanism of Light-Induced Signaling. *ChemBioChem* **2006**, *7*, 187−193.

DOI: 10.1021/acs.jctc.5b00197
J. Chem. Theory Comput. 2015, 11, 3878−3894