Article

Influence of the Environment on Shaping the Absorption of Monomeric Infrared Fluorescent Proteins

Sivasudhan Rathnachalam, Maximilian F. S. J. Menger, and Shirin Faraji*

Cite This: *J. Phys. Chem. B* 2021, 125, 2231–2240

Read Online

ACCESS | Metrics & More | Article Recommendations | SI Supporting Information

ABSTRACT: Infrared fluorescent proteins (iRFPs) are potential candidates for deep-tissue in vivo imaging. Here, we provide molecular-level insights into the role of the protein environment in the structural stability of the chromophore within the protein binding pocket through the flexible hydrogen-bonding network using molecular dynamics simulation. Furthermore, we present systematic excited-state analysis to characterize the nature of the first two excited states and the role of the environment in shaping the nature of the chromophore's excited states within the hybrid quantum mechanics/molecular mechanics framework. Our results reveal that the environment red-shifts the absorption of the chromophore by about 0.32 eV compared to the isolated counterpart, and besides the structural stability, the protein environment does not alter the nature of the excited state of the chromophore significantly. Our study contributes to the fundamental understanding of the excited-state processes of iRFPs in a complex environment and provides a design principle for developing iRFPs with desired spectral properties.
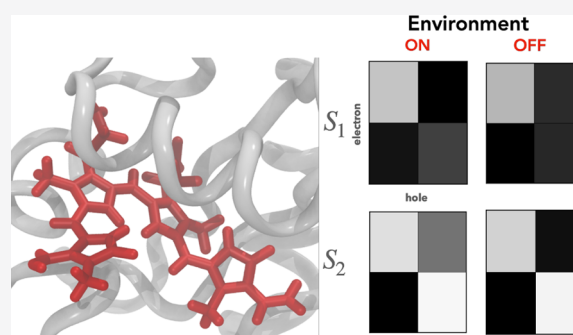
1. INTRODUCTION

Ever since the successful derivation of genetically encoded fluorescent protein (FP), it became a significant probe in the field of bio-imaging.[1] The FPs are widely used in studying protein dynamics inside the cellular environment[2] and as a marker for gene expression.[3] The quantum yield of fluorescence was not satisfactory in the wild-type (wt) GFPs for fluorescence imaging due to low molecular extinction coefficients and part of the absorption band in the ultraviolet region.[4] Spectroscopic analyses revealed that the chemical structure, the charge state, and the immediate environment of chromophores are crucial in determining their spectral property.[5] For instance, an anionic chromophore has red-shifted absorption/emission relative to the neutral form in GFPs.[6] Over the past decades, the initially discovered GFPs in *Aequorea victoria* (jellyfish) have undergone several steps of structural engineering of chromophores and site-directed mutations of the protein, leading to expansion of the fluorescence color palette from bimodal absorption peaks (395 nm and 475 nm) in wt-GFPs to the entire visible range of the electromagnetic spectrum.[7,8] Red fluorescent proteins (RFP) became popular in bio-imaging of mammalian cells due to the low energy of the red light, higher penetrability in thick tissues, and reduced autofluorescence, thus resulting in a high signal-to-noise ratio.[9−11] However, the fluorescence imaging of mamallian tissues still poses a challenge as the absorption of heme, melanin, and water is also in the same region of RFPs. Recently developed near-infrared FPs (iRFPs) from bacterial phytochrome photoreceptors (BphPs)[1] offer a new twist to the conventional FPs and have an optical window in far-red or near infrared (NIR) (600−900 nm). Additionally, iRFPs have a heme-based chromophore that is endogenous to mammalian cells and could auto-catalytically bind to a protein.[12] Therefore, the iRFPs are widely studied for applications such as optogenetics,[13] detecting and modulating protein−protein interactions,[13,14] and bio-imaging of mammalian tissues.[15]

The BphPs serves as a template for NIR FPs, belonging to the phytochrome family of photoreceptors and found in various living organisms such as plants, algae, and bacteria.[17] The BphPs incorporate a heme-derived biliverdin IXα (BV), a linear tetrapyrrole molecule, as the chromophore.[18] Typically, the photosensory module of BphPs has three domains responsible for light sensing, namely, Per-ARNT-Sim (PAS), cGMP phosphodiesterase/adenylate cyclase/FhlA (GAF), and phytochrome-specific (PHY).[13,19] The photoabsorption of BV triggers chain of structural transformation in the photosensory module.[20] Upon photoabsorption, the BphPs could switch between P_r (Z isomer) and P_{fr} (E isomer) via rotation around the double bond between C_{15}=C_{16} (see Figure 1a), thereby
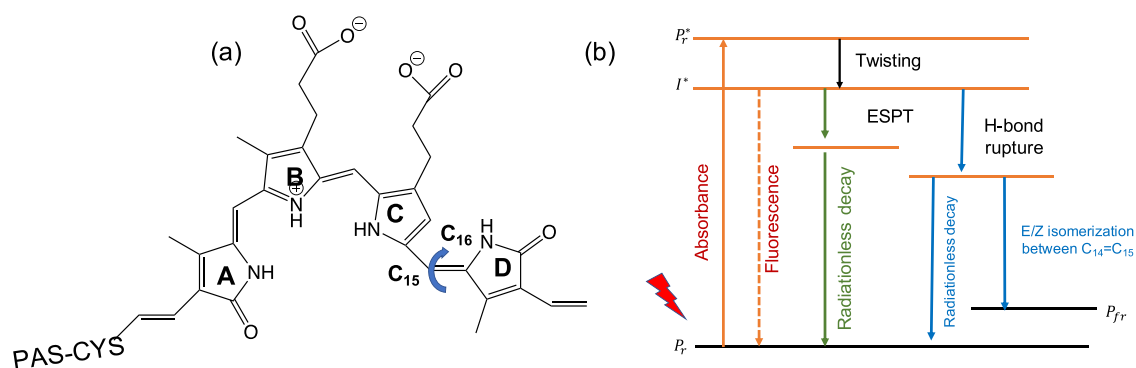
Figure 1. (a) Chromophore biliveridin IXα (BV). BV is covalently bound to the protein via sulfur of the Cys residue of the PAS domain. The pyrrole rings of BV are labeled as A, B, C, and D. The blue arrow indicates the rotation around $C_{15}=C_{16}$, resulting in photoisomerization to the P_{fr} state. (b) Various excited-state decay channels of BV in bacteriophytochromes (dimeric form) such as excited-state proton transfer (ESPT), photoisomerization, and fluorescence are shown.[16]

reducing the quantum yield of fluorescence in wt-BphPs. In iRFPs, photoisomerization and excited-state proton transfer (ESPT) were identified to be the significant competing excited-state decay channels to the desired fluorescence (Figure 1b).[16]

A recent computational study on a (monomeric) DrBphP system showed that nonclassical hydrogen bonding such as the OH-π interaction of Tyr residue with ring D of BV along with hydrogen-bonding partners in binding pocket is required for the conformational stability of the chromophore.[21] The molecular dynamics (MD) combined with X-ray crystallography study concluded that rigidity of the rings is important for enhanced fluorescence while the presence of water−BV interaction diminishes fluorescence by reducing the planarity of BV.[22] A computational study on DrBphP revealed that the absorption spectrum of BV was dominated by the contributions of a protonated BV chromophore with traces of deprotonated pyrrole D ring.[23] Furthermore, pyrrole ring C was identified to be more susceptible to ESPT.[24]

Engineering the protein environment of the wt-BphPs by removing the PHY domain has resulted in restraining the pyrrole ring D, thereby blocking the photoisomerization decay channel of BV[25] but yet the ESPT channel quenching the fluorescence.[26] In engineered proteins derived from an RBphP1 template, the removal of the PHY domain and with key point mutations such as Asp216Ala resulted in preventing hydrogen-bonding interaction with BV, thereby effectively blocking the ESPT channel.[18,27] Recently, the crystal structure of the monomeric form of iRFPs (miRFPs) became available.[12] The miRFPs (derived from BphPs) are composed of only PAS and GAF domains, excluding the PHY domain, and the chromophore BV is covalently bound to the Cys residue of the PAS domain.[12,28,29] These engineered miRFPs have improved the fluorescence quantum yield 5-fold compared to the wt-BphPs.[15]

GFP-derived proteins constrain the chromophore within the binding through the β-barrel structure pocket.[5] Structural stability of the chromophore is required in order to avoid the undesired nonradiative heat dissipation.[30] For this reason, numerous mutants of the chromophore derived from wt-GFPs preserved the native β-barrel structure. However, the role of the environment in the miRFPs remains unclear, as the theoretical understanding of the excited-state processes in bulky systems (such as iRFPs) remains sparse due to the size of the system and the challenges involved in describing the

excited-state processes with a reliable electronic-structure method.[31] It was shown that *ab initio* methods such as XMCQDPT2 could compute the absorption spectra of BV in agreement with the experimental spectra.[23] In RFPs, it was shown computationally that slight variation in hydrogen-bonding interactions of the chromophore with the immediate environment could shift the absorption band of the fluorescent proteins.[32] The extended Stokes shift in mPlum protein (Far RFPs) was explained with the stability of excited state due to a flexible hydrogen-bonding network of the protein environment.[33] The interaction of pyrrole ring D of BV with the conserved Tyr142 residue from the protein environment controls the photoisomerization process in the billin-binding phytochromes.[34,35] The influence of hydrogen-bonding interactions of the chromophore with the environment in BphPs (dimeric) was found to be insignificant in computing the absorption spectra of BV.[36] To the best of our knowledge, the role of the environment in the stability of the miRFP chromophore within the binding pocket as well as its spectral properties remained unexplored. Therefore, the key objective is to understand the influence of the environment on shaping the excited-state character of BV in miRFP. Our findings could motivate future studies toward controlling the excited-state dynamics of miRFPs and suggest clues for rational mutagenesis of miRFPs.

The paper is structured as follows: Section 2 details the quantities used in the chromophore's excited-state analysis. Section 3 describes the theoretical approaches and the computational details. Following the computational section, the ground-state MD results are analyzed to unravel the interactions that stabilize the chromophore in the binding pocket. Subsequently, the influence of the environment is analyzed through excited-state analysis across the MD trajectory. Concluding remarks are presented in Section 5.

2. THEORETICAL METHODS

In the photoabsorption process, the light absorbed by a molecule results in the formation of a quasi-particle called exciton. To unravel the nature of excited-state transitions, one-particle transition density matrix (1TDM)-based exciton analyses were performed. The exciton wavefunction can be written as a product function of the 1TDM of ground state (Φ^0) and excited state (Φ^I). Matrix representation of the 1TDM in an orbital basis set $\{\chi_\eta\}$ can be represented as

$$\gamma^{0I}(r_h, r_e) = \sum_{\mu\nu} D_{\mu\nu}^{0I} \chi_\mu(r_h)\chi_\nu(r_e) \tag{1}$$

where $\chi_\mu(r_h)$ and $\chi_\nu(r_e)$ represents the μth and ν^{th} orbitals where the electron−hole and the excited electron are located, respectively. The 1TDM element $D_{\mu\nu}^{0I}$ is expressed in terms of one-particle creation $(\hat{a}_\mu^\dagger)$ and annihilation $(\hat{a}_\nu)$ operators as $D_{\mu\nu}^{0I} = \langle \psi^0 \mid \hat{a}_\mu^\dagger \hat{a}_\nu \mid \psi^I \rangle$. The quantity γ^{0I} can be used in computing the exciton descriptors that allow identifying and characterizing the nature of the excited states of a molecule. The excited-state descriptors used are briefly described here. For elaborate discussions on excited-state descriptors, the reader is referred to refs 38 and 39.

Partially integrating the square of the 1TDM while restricting the hole to a fragment A and the electron to a fragment B is represented as

$$\Omega_{AB} = \int_A \int_B \gamma^{0I}(r_h, r_e)^2 dr_e dr_h \tag{2}$$

where Ω_{AB} is the charge-transfer (CT) numbers. The Ω_{AB} represents the probability of the electron density transferred between the fragments during the excitation. For the cases where more than two fragments are considered, the Ω_{AB} can be analyzed using the so-called electron−hole correlation plots. As the system is partitioned into n fragments, the Ω_{AB} computed for different fragments is represented as a pseudo-color matrix plot with size $n \times n$ (fragments). The diagonal elements of the matrix correspond to local excitation (LE), and the off-diagonal elements are characterized as CT transitions. The sum of the off-diagonal contributions gives the total CT number (ω_{CT}) and is defined as

$$\omega_{CT} = \frac{1}{\Omega} \sum_{A,B \neq A} \Omega_{AB} \tag{3}$$

where Ω is the norm of the 1TDM:

$$\Omega = \langle \gamma^{0I} | \gamma^{0I} \rangle \tag{4}$$

The ω_{CT} values range from 0 to 1 where 0 corresponds to local excitation and one correspond to a completely separated CT state.

Unfortunately, theoretical identification of a CT state is not straightforward and it is important to distinguish between two states: (1) CT states with permanent shift of electron density and (2) charge-resonance (CR) in which the net CT is zero. CT and CR states can be distinguished by measuring the separation of electron and hole centroids and is expressed as

$$d_{h\to e} = |\langle \vec{x}_e - \vec{x}_h \rangle_{exc}| \tag{5}$$

The difference between the CR and LE states can be understood by taking the spatial distribution of electron and hole into account through the so-called exciton size, d_{exc}, which is expressed as the root-mean-square separation of electron−hole

$$d_{exc} = \sqrt{\langle |\vec{x}_e - \vec{x}_h|^2 \rangle_{exc}} \tag{6}$$

The d_{exc} values differ significantly for LE and CR states. Last, to describe the statistical correlation between an electron and hole, we use the electron−hole correlation coefficient, R_{eh}, defined as

$$R_{eh} = \frac{COV(r_h, r_e)}{\sigma_h \sigma_e}, \tag{7}$$

where $COV(r_h, r_e) = \langle \vec{x}_h \cdot \vec{x}_e \rangle_{exc} - \langle \vec{x}_h \rangle_{exc} \cdot \langle \vec{x}_e \rangle_{exc}$, and it describes spatial relation between an electron and hole. R_{eh} ranges from −1 to 1, values greater than zero would mean a bound-exciton, negative values imply that the electron and hole avoid each other in space, and zero would mean lack of spatial correlation between the electron and hole.

To visually represent the electron hole localization across the molecule, we used natural transition orbitals (NTOs). The single-value decomposition of the 1TDM results in NTOs and gives a compact state-specific description of the excited state of interest.[37]

3. COMPUTATIONAL DETAILS

The crystal structure for the representative miRFPs was obtained from the protein data bank database (PDB ID: 5VIK).[12] The structures of the BV chromophore and the protein were checked for missing H atoms and added appropriately (at neutral pH) such that the net charge of the chromophore was −1.0 and that of protein was −3.0. The molecular mechanics simulations were performed using GROMACS[39] with the parameters obtained from the AMBER protein force field (ff14SB) for the standard residues of the protein. For the non-standard BV residue, generalized AMBER force field (GAFF)[40] was used and the parameters are given in the Supporting Information (Tables S5−S7), while the atoms in Cys part used the standard definition of atom type in Amber protein force field along with the fitting parameters from Rao et al.[41]

The protein along with the chromophore was placed at the center of a cubical box of size 102.7 Å, separated by 20 Å from the walls, and solvated with explicit TIP3P waters.[42] The system was neutralized by adding 4 Na$^+$ ions. The LINCS[43] algorithm was employed to constrain the C−H bond length. Additionally, periodic boundary conditions were used with the particle mesh Ewald scheme[44] and switching function for van der Waals interaction.[45] The system was thoroughly equilibrated in several steps. The externally added H atoms to the proteins were allowed to relax by setting very high positional restraint (fc = 1000 kJ mol^{-1}nm^{-2}) on the protein and chromophore heavy atoms. First, to minimize the steric clashes within the protein and protein−water, energy minimization was performed. Second, the minimized structure was thermally equilibrated to 300 K through constant particle, volume, and temperature (NVT) ensemble with the same positional restraints on the heavy atoms for 100 ps. Third, the constant particle, temperature, and pressure ensemble (NPT) using a velocity-rescaling thermostat[46,47] and an isotropic pressure coupling[48] with 2 ps time constant was achieved using a Parinello-Rahman barostat.[49] NPT equilibration was carried out for 100 ps such that the density of the system is stabilized at 1 atm and 300 K. To ensure gradual release of restraints, we repeat the aforementioned steps with varying harmonic restraints at 25, 10, and 5 kJ mol^{-1}nm^{-2} values on the protein and chromophore heavy atoms. The equilibrated structure was obtained through an unconstrained energy minimization and thermal equilibration (NVT) for 1 ns. Finally, the starting geometry of the system was obtained after 1 ns of NPT equilibriation. This procedure helps in gradual equilibration of heavy atoms along with H atoms, solvents, and ions. We obtained the conformational sampling of the system by

carrying out a *NVT*-production MD using a 2 fs time step and without any restraints for 250 ns. The conformational space of the chromophore in the trajectory was clustered using the GROMOS algorithm.[50] The hydrogen-bonding analysis of the MD trajectory was performed using the cutoff distance of the donor and the acceptor at 3.0 Å and the angle between them at 30°.

The BV-I chromophore structure obtained from the crystal structure (PDB ID: 5VIK) served as the initial structure for optimization using the ωB97X-D with cc-pVDZ basis set including Grimme's dispersion correction.[51] Hessian calculation was performed to confirm the minimum structure. The computation of vertical excitation energies (V_{EE}) and the subsequent excited-state analysis were obtained using time-dependent density functional theory (TD-DFT); we picked the ωB97X-D/cc-pVDZ level of theory as the CT character of various molecules was well recovered using range-separated functionals.[52−54] To study the isolated chromophore, we considered two models (see Figure 2): (i) BV without carboxylic tails (BV-I) and (ii) BV with carboxylic tails protonated (BV-II).
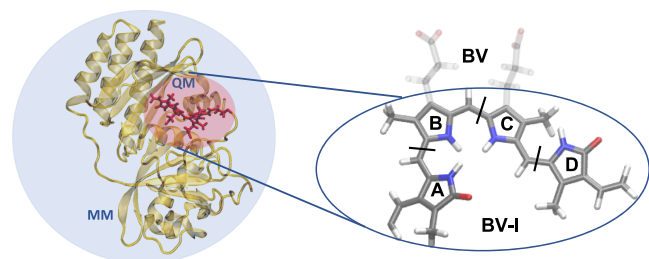
Figure 2. Schematic representation of the QM region (BV-I): the chromophore (without carboxylic tails). The rest of the system including carboxylic tails of BV, protein, water, and ions was embedded as MM point charges. BV with carboxylic tail protonated is denoted as BV-II. The black lines in the chromophore represents the fragmentation scheme followed for the excited-state analysis.

To incorporate the influence of the environment onto the excited-state character of the chromophore, we employed the quantum mechanics/molecular mechanics (QM/MM) approach. The QM region consists of chromophore BV (according to the PDB structure[12]), and the chromophore was cut off at the covalent linkage of BV and Cys residue and the carboxylic tails from pyrrole rings B and C (see Figure 2). The QM region includes the chromophore without carboxylic tails (BV-I). The rest of the system including the carboxylic

tails, protein, water, and counterions was treated as the MM region. In a hybrid QM/MM scheme (Figure 2), the interaction was described as the electrostatic embedding of MM point charges that is added as one-electron operator to the QM Hamiltonian. The structures for QM/MM were taken from the MD trajectory at every 1 ns. Link atoms are used to saturate free valencies of the boundary QM atoms when the QM and MM subsystem are separated. Overpolarization[55] is prevented by using the charge distribution scheme, in which the charge on the frontier MM atom (S of the Cys residue) is deleted and equally distributed over the other covalently bound MM atoms (MM host atoms). The advantage of this approach is that the charge of the overall system is preserved. All electronic structure calculations were done using Q-Chem 5.2.[56] Post-processing of the excited-state analysis was done using a TheoDORE program.[57] The graphical visualization of the trajectory was done using a VMD program.[58]

4. RESULTS AND DISCUSSION

4.1. Interaction of the Chromophore with Its Immediate Environment. In order to understand the structural stability of the chromophore in the binding pocket of the protein, we analyzed the 250 ns trajectory obtained from the *NVT* production run. The root-mean-square deviation (RMSD) plot (Figure 3a) suggests that the structure of BV fluctuates predominantly between two RMSD values (around ~0.8 and ~1 Å) in an interval of 0−120 ns, after which the RMSD of BV fluctuates around 1.0 Å.

To further understand the structural fluctuations of the chromophore, we performed clustering analysis based on RMSD. It became apparent that the chromophore structure fluctuates between two structures predominantly (Figure S2), which differ by the "floppy" dihedral that includes ring A (A1⋯A4), ring D (D1⋯D4), and carboxylic tails on the ring C (T21⋯T24) and on the ring B (T11⋯T14) (see Figure 3b).

In order to characterize the hydrogen-bond network around the chromophore and identify the residues that contribute to the network, we have performed hydrogen-bond analysis over 250 ns simulation. Hydrogen-bond occupancy values, representing the percentage of times the residue−chromophore (atom) pair was at hydrogen-bonding criteria, are presented in Table 1. It is evident that the residues Arg230 and Arg198 have the highest occupancy numbers, i.e., more than 90% of the simulation time. A closer inspection reveals that these two basic amino acids are hydrogen-bonded to the charged carboxylic tails of BV, stabilizing them within the binding pocket (see Figure 4). Additionally, two polar amino
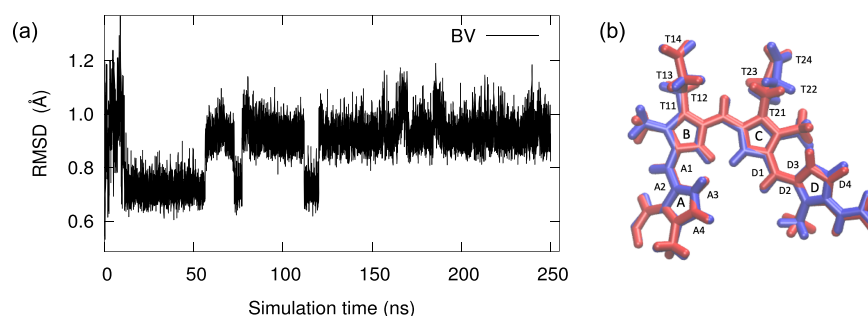

Figure 3. (a) RMSD of the chromophore structure across the trajectory. (b) Right panel shows the averaged structure of the most dominant clusters, cluster-1 (red) and cluster-2 (blue) of the 250 ns trajectory, superimposed. Atoms in the selected dihedral are labeled (only heavy atoms) for structural analysis of clusters.

Table 1. Hydrogen-Bonding Pairs of Chromophore Atoms and the Neighboring Amino Acids[a]

donor	acceptor	distance [Å]	occupancy (%)
Arg230 (H1)	BV (O1)	1.84 (0.15)	94.8
Arg230 (H2)	BV (O2)	1.87 (0.18)	90.6
Ser233 (H3)	BV (O2)	2.08 (0.47)	59.95
Arg198 (H4)	BV (O3)	2.15 (0.23)	95.11
Tyr192 (H5)	BV (O1)	2.12 (0.53)	68.46
Arg198 (H6)	BV (O4)	2.15 (0.23)	41.22
Arg198 (H7)	BV (O4)	1.85 (0.13)	39.22
BV (H8)	Leu183 (O5)	2.17 (0.21)	34.43
His266 (H9)	BV (O6)	2.38 (0.47)	33.41
BV (H10)	Leu183 (O7)	2.15 (0.21)	39.3

[a]Average bond distance of a hydrogen-bond pair (donor being the moiety donating hydrogen). Standard deviations are given in parentheses, and the occupancy values were computed across the 250 ns trajectory.
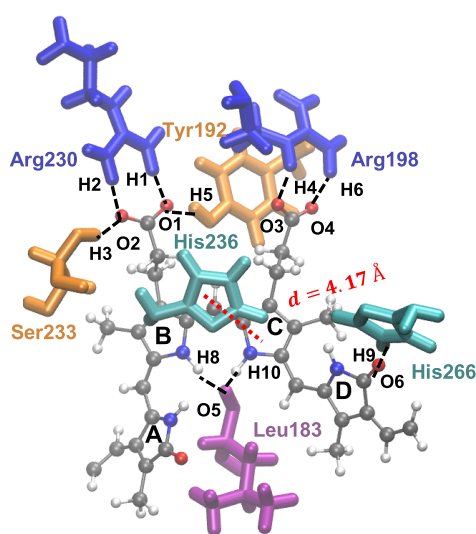
Figure 4. Schematic representation of hydrogen-bonding interactions found between BV and the protein during MD simulation. BV is bound by electrostatic interactions of basic amino acids (Arg198 and Arg230), polar amino acids (Ser230, Tyr192, and His266), and a nonpolar amino acid (Leu183). $\pi - \pi$ stacking interaction of His236 with BV (ring C). The distance of separation, d [Å], between the center of masses of the five-membered rings of His236 and C-ring of the chromophore is presented by a red dotted line.

acids Ser23 and Tyr192 residues also participate in hydrogen-bonding interaction with the carboxylic tails of the BV chromophore (see Figure 4). Moreover, the hydrogen-bond interaction of His266, with 33% occupancy, to BV (ring D) would possibly explain the tilt (or distortion) in the ring D of BV (see Figure 3b).

It must be noted that in the conventional iRFPs, including the dimeric form of the miRFP studied in this work, the polar Asp residue is hydrogen-bonded to the rings B and C of BV, while in the miRFP, Asp is replaced by a nonpolar residue, Leu183.[12] Our results shows that Leu183 also remains hydrogen-bonded to the rings B and C throughout the simulation, although with lower occupancy values (Table 1). Last but not the least, four water molecules contribute to the hydrogen-bonding network around the chromophore throughout the simulation (see Figure S3), including the encapsulated water between the rings A, B, and C of BV.

Additionally, our results demonstrates that the $\pi - \pi$ stacking interaction of His236 with the ring C of BV present in the crystal structure remained preserved throughout the simulation (Figure 4). The values of separation of ring centroids (d) averaged over the 250 ns trajectory is 4.17 Å (with a standard deviation of 0.36), suggesting that the ring centroids fall very close to the corresponding value in the crystal structure (4.1 Å).[12]

4.2. Vertical Excitation of the Chromophore. *4.2.1. Isolated Chromophore.* The hydrogen-bonding interaction holds carboxylic tails of the chromophore with polar and basic residues from the protein environment (see Figure 4). Therefore, we study the BV chromophore in two structural forms, namely, BV-I and BV-II. In the literature, the ωB97X-D functional in the TD-DFT/Tamm−Dancoff approximation framework was successful in determining the absorption behavior of the BV chromophore.[36] Additionally, it was also shown that the exciton descriptors are better described using range-seperated functional CAM-B3LYP.[52,53] It is to be noted that ωB97X-D has similar percentage of Hartree−Fock exchange to CAM-B3LYP, and since the structures were also optimized using ωB97X-D, we use the former for studying the excited-state properties of the chromophore. We also find that excited-state descriptors of isolated BV-I and BV-II computed using CAM-B3LYP functional (Table S2) are similar to the ones obtained using ωB97X-D (Table 2).

In Table 2, V_{EE} values of the first two bright states of BV-I and BV-II appear to be quite similar (difference of ~0.03 eV for the S_1 state and ~0.1 eV for the S_2 state). For the 1-TDM-based excited-state analysis, four fragments were considered; the fragmentation scheme (Figure 2) and their corresponding electron−hole correlation plots are given in Figure S4. First, we discuss the nature of the S_1 state. The ω_{CT} values are around 0.6, indicating that the nature of excited state is impure, i.e., mixed LE and CT characters. Both $d_{h \to e}$ and d_{exc} values are similar for both BV-I and BV-II models (difference of 0.1 Å), indicating that the natures of the excited states are similar, i.e., the LE character. The positive but small value of R_{eh} (~0.18) suggests lesser Coloumbic attraction of the electron−hole pair. Looking at the S_2 state, we observe that the $d_{h \to e}$ in BV-I is higher by 0.6 Å, suggesting that the mix of CT character is slightly higher as compared to the predominant LE nature of S_2 of BV-II. The d_{exc} values of BV-II is slightly higher than those of BV-I, suggesting larger delocalization of the exciton.

Table 2. Excited-State Descriptors of the Two Bright States of the Isolated Chromophore in Two Model Forms, BV-I and BV-II (See Text for Explanation), and the Fragmentation Scheme Described in Figure S4

system	state	V_{EE} [eV] (f)	ω_{CT}	$d_{h \to e}$	d_{exc}	R_{eh}	NTO (weight)
BV-I	S_1	2.29 (1.0)	0.664	0.563	5.081	0.188	$h_1 \to e_1$ (0.93)
	S_2	4.04 (1.1)	0.601	1.380	5.064	0.317	$h_1 \to e_2$ (0.61)
BV-II	S_1	2.32 (0.9)	0.654	0.671	5.030	0.184	$h_1 \to e_1$ (0.91)
	S_2	4.11 (1.2)	0.602	0.784	5.235	0.284	$h_1 \to e_2$ (0.64)

Comparable values of R_{eh} are obtained for both the models of the chromophore at the S_2 state.

The electron–hole correlation plots (see Figure S4) show that the exciton in both the S_1 and S_2 states are delocalized over the entire molecule. Given the similarities of the descriptors of the first two bright states of BV-I and BV-II, the effect of carboxylic tails of the chromophore is insignificant in describing the nature of the excited states. Therefore, in the following, we restrict ourselves with calculations involving BV-I as the QM region.

4.2.2. Chromophore with the Environment. To elucidate the role of the environment in shaping the chromophore's absorption, we considered 250 structures from the MD trajectory. Table 3 shows the computed excitation energies of various model systems as well as the experimental values.

Table 3. Vertical Excitation Energies, V_{EE} in eV, and Oscillator Strength (f) in Parentheses[e]

system	MM	S_1	S_2
BV-I[a]		2.29 (1.0)	4.04 (1.11)
BV-I[b,c]	all	1.97 (1.31)	3.63 (1.39)
BV-I[c]		2.06 (1.4)	3.67 (1.43)
experimental[d]	all	1.83	3.6

[a]Single optimized structure of isolated BV-I. [b,c]Average energies computed using 250 structures obtained from the MD trajectory. [d]Ref 12. [e]V_{EE} values were computed at the ωB97X-D/cc-pVDZ level of theory.

The difference in the average (over 250 structures) excitation energy of the chromophore with the electrostatic embedding of MM point charges (BV-I[b]) and the isolated (BV-I[c]) appear to be minimal (0.09 eV). Additionally, we performed excited-state analysis using the representative structure of BV-I taken from 250 ns MD trajectory, with and without the MM point charges. We chose the average structure of the most predominant cluster from the 250 ns trajectory as a representative structure for our analysis.

We begin by comparing the S_1 state of the isolated BV-I and BV-I with the inclusion of the MM point charges (see Table 4). The V_{EE} values are quite similar, with a difference of ∼0.1 eV. The $d_{h \rightarrow e}$ values are slightly larger for BV-I with the MM point charges, suggesting that the exciton is slightly more delocalized while the d_{exc} values are comparable (∼5.6 Å). Again, a slightly higher R_{eh} value suggests that MM point charges around BV-I could help in larger exciton binding due to Coulombic attraction. It must be noted that the ω_{CT} value strongly depends on the fragmentation scheme. For example,

here we observe higher ω_{CT} values with an increase in the number of fragments ($n = 2,3,$ and 4) considered. Interestingly, the relative difference between the ω_{CT} values of BV-I with and without MM point charges remain insignificant for all the fragmentation schemes. The S_2 state has similar V_{EE} both in isolated BV-I and BV-I in the presence of the MM point charges (∼3.6 eV), further implying that the influence of the environment in the S_2 state character of BV-I is minimal. In sum, the electrostatic embedding of the MM charges has an insignificant effect in both S_1 and S_2 excited-state descriptors.

To further understand the nature of the excited-states involved, we analyze the electron–hole correlation plots with different fragmentation schemes, i.e., varying number of fragments (Figure 5). First, we discuss the results of BV-I with two fragments (A and B): the S_1 state appears to have similar character with and without the MM charges, where the exciton appears to be delocalized across BV-I. For the S_2 state, isolated BV-I has slightly more LE in fragment A as compared to BV-I embedded with MM point charges that has LE on both fragments A and B. The S_2 state has minor CT contribution between fragment B → A. Therefore, it appears that the S_2 state has a predominant LE nature.

In the case of three fragments (A, B, and C), the S_1 state has a dominant LE character that involves fragment B, i.e., pyrrole rings B and C of BV-I. However, there are numerous minor CT contributions additionally between all the three fragments (except no CT transitions from fragment A → C and vice versa). Therefore, having LE at fragment B and other CT transitions between the fragments leaves the exciton delocalized across the molecule, making S_1 a delocalized excited state. The S_2 state of BV-I, with the presence of the environment, has slightly larger localization of exciton in all the three fragments, while in the isolated BV, the LE contribution of fragment A becomes more significant than the others. The CT transitions across the fragments become more pronounced with the inclusion of point charges.

On fragmentation of BV-I into fragments (A, B, C, and D), we once again see that the role of electrostatic interactions of the immediate environment in perturbing the nature of the excited states remains minimal. All four fragments of the S_1 state seem to have LE transitions, and even the CT transitions from fragments A → B, C → B, and C → D remain similar for both BV-I with and without the environment. For the S_2 state, the isolated BV-I has dominant LE in fragment A with a set of minor CT transitions, and the trends remain similar for BV-I with the inclusion of the MM point charges.

From the discussion above, it is apparent that the electron–hole correlation plots with various fragmentation schemes lead

Table 4. Excited-State Analysis of the Representative Structure of BV-I across the 250 ns Trajectory[a]

type	state	V_{EE} [eV] (f)	$d_{h \rightarrow e}$ [Å]	d_{exc} [Å]	ω_{CT} $n = 2$	$n = 3$	$n = 4$	R_{eh}	NTO (weight)
QM/MM	S_1	1.93 (1.35)	1.063	5.679	0.371	0.501	0.647	0.246	$h_1 \rightarrow e_1$ (0.89)
	S_2	3.60 (1.12)	0.722	4.728	0.166	0.476	0.526	0.478	$h_1 \rightarrow e_2$ (0.28)
									$h_2 \rightarrow e_1$ (0.20)
isolated	S_1	2.02 (1.45)	0.728	5.649	0.405	0.501	0.668	0.205	$h_1 \rightarrow e_1$ (0.90)
	S_2	3.62 (0.98)	0.621	4.716	0.152	0.409	0.446	0.487	$h_1 \rightarrow e_2$ (0.25)
									$h_2 \rightarrow e_1$ (0.24)

[a]The structure was studied with electrostatic embedding of MM point charges around BV-I (QM/MM) and as isolated BV-I. The first two bright states of the cluster were computed at the ωB97X-D/cc-pVDZ level of theory, with various fragment numbers, n (see Figure 5 for the fragmentation schemes).
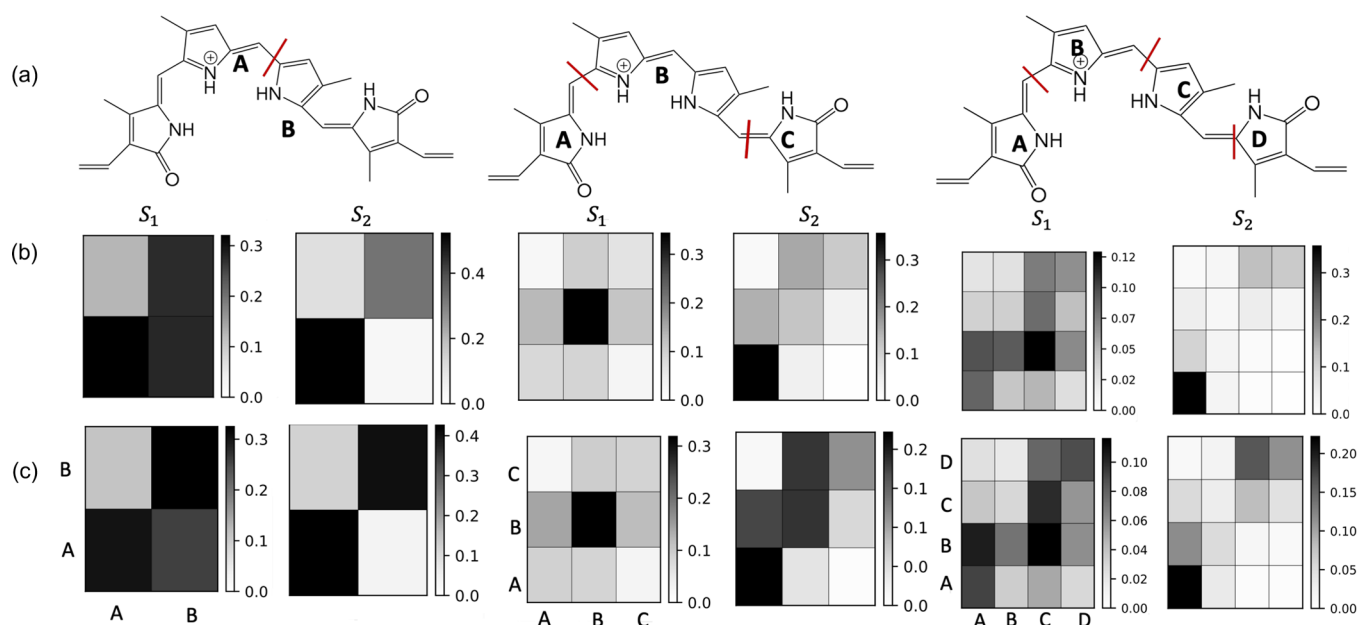
Figure 5. Electron−hole correlation plots of the representative structure of the MD trajectory with various fragmentation schemes. (a) Fragmentation scheme employed, (b) BV-I (isolated), and (c) BV-I with MM point charges.
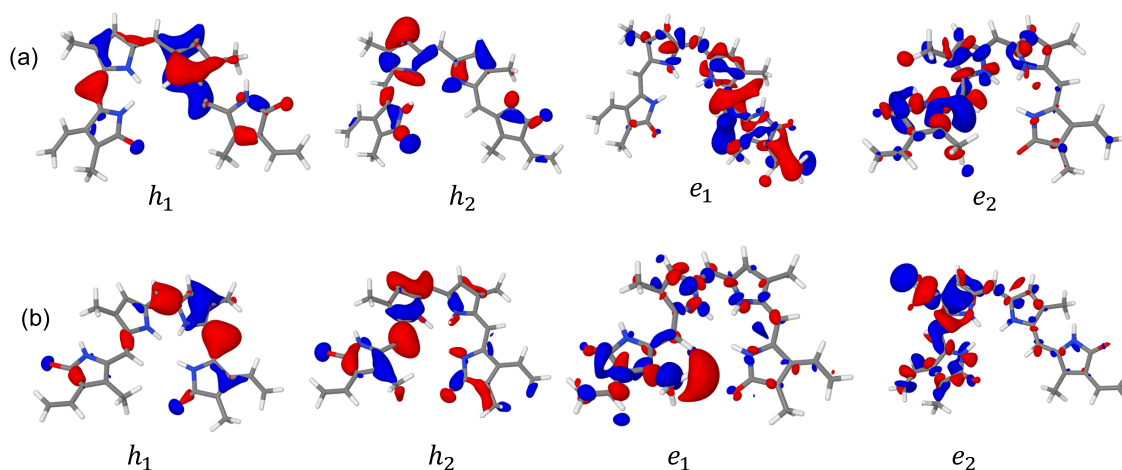
Figure 6. State-averaged NTOs of the representative structure of BV-I from the MD trajectory: (a) isolated and (b) with electrostatic embedding of MM charges using an isovalue = 0.03.

to different interpretations. However, the key finding is that in all the three aforementioned fragmentation schemes, the exciton is spread across all the fragments for both S_1 and S_2 states, leaving the nature of the excited states a delocalized one. It is worth mentioning that a recent study,[59] based on exciton descriptors that are independent of the fragmentation scheme, has reported that within the TD-DFT framework, using the correct amount of nonlocal orbital exchange functional, the d_{exc} has a universal value of amount 7 Å for π-conjugated polymers. As it is evident from Table 4, the d_{exc} values obtained for our system are around 4−5.7 Å, which are much smaller that 7 Å, suggesting that BV-I is dissimilar to typical extended π-conjugated systems. Interestingly, the nonplanar BV-I with repeating units of five-membered heteroaromatic cycle is structurally similar to the hexa-thiophenes, which were found to have a delocalized excited state with a d_{exc} value of 5 Å[53] comparable to the d_{exc} values obtained for BV-I.

NTO transitions and their corresponding weights are given in Table 4. The state-averaged NTOs (SA-NTOs) are

designated as h_i and e_i orbitals corresponding to the ith orbital of hole and electron. SA-NTOs (Figure 6) reveal that (i) 90% of the S_1 state is represented by the transition $h_1 \rightarrow e_1$, which is $\pi \rightarrow \pi^*$. (ii) S_2 states are composed of majorly two transitions $h_1 \rightarrow e_2$ and $h_2 \rightarrow e_1$. The SA-NTOs are in line with the electron−hole correlation plots, suggesting delocalized S_1 and S_2 states. It must be noted that BV-I in Table 2 is QM-optimized while in Table 4, it is a single-point calculation of the representative structure obtained from a 250 ns MD trajectory and therefore their corresponding NTO transitions differ.

To make the analysis statistically significant, we analyze the distributions of the excited-state descriptors (with and without the MM point charges) for 250 structures obtained from the MD trajectory.

4.2.3. Excited-State Descriptors of BV-I across the Trajectory. To statistically characterize the excited-state character of the first two bright states of BV-I, we followed

the exciton descriptors' distributions across the MD trajectory with four fragments, as described in Figure 5.

Table 5 shows the average and the standard deviation values of the excited-state descriptors. In the case of the S_1 state, the

Table 5. Excited-State Descriptors of BV-I Averaged over 250 Structures with Electrostatic Embedding of MM Point Charges around BV-I (QM/MM) and as Isolated BV-I[a]

type	state	$d_{h \to e}$ [Å]	d_{exc} [Å]	R_{eh}
QM/MM	S_1	1.388 (0.20)	5.745 (0.08)	0.249 (0.01)
	S_2	0.676 (0.43)	5.115 (0.28)	0.389 (0.06)
isolated	S_1	0.968 (0.13)	5.669 (0.06)	0.206 (0.01)
	S_2	1.051 (0.46)	5.068 (0.39)	0.415 (0.07)

[a]The standard deviation of the descriptors are given in parentheses. The first two bright states of the cluster were computed at the ωB97X-D/cc-pVDZ level of theory.

d_{exc} values are in a range of 5.6−5.8 Å, which is larger than the fragments considered (∼4.3−4.5 Å, see Figure 5). Similar to d_{exc}, $d_{h \to e}$ values are consistently high for the S_1 state, suggesting a slightly larger delocalization of the exciton in the presence of MM point charges. The R_{eh} values are also similar (with a difference of 0.05 units), suggesting that the role of the electrostatic interactions of MM charges in stabilizing the exciton binding is minimal too.

In contrast to S_1, for the S_2 state, the average values of $d_{h \to e}$ of BV-I in the presence of MM point charges are lower, implying that the exciton is slightly more localized (see Table 5). The d_{exc} values of BV-I are in the presence of MM point charges, and the values are comparable to isolated BV-I. The R_{eh} is in a range of 0.38−0.41, slightly larger compared to the S_1 state that suggests that the exciton is bound and stabilized by Coulombic interactions. The excited-state descriptor values obtained across the trajectory are in line with the analysis of the representative structure of BV-I (Table 4). Therefore, we could conclude that the nature of first two bright states remains delocalized throughout the trajectory. Additionally, the plots of Ω-matrix and diagonal (LE) and off-diagonal (CT) elements across the trajectory for BV-I are given in the Supporting Information. The nature of the plots (see Figure S7−S10) corroborates with our previous analysis, namely, (1) the effect of electrostatic interaction stabilizing the exciton is minimal and (2) the exciton is spread across all four fragments, speaking in favor of the fact that both S_1 and S_2 are delocalized excitations.

5. CONCLUSIONS

The role of the environment in influencing the absorption of the BV chromophore was studied using 1TDM-based exciton analysis within a hybrid QM/MM framework. The MD results suggest that although the hydrogen-bond network around the chromophore is sufficiently flexible, many of the crystal contacts of the chromophore with the proteins are well-preserved. The dynamic hydrogen-bonding interactions around the chromophore ensure the structural stability of the chromophore within the protein binding pocket. In particular, basic amino acids such as Arg230 and Arg198 and polar amino acids Tyr192 and Ser233 along with the encapsulated water play a vital role in the structural stability of the chromophore. Inclusion of the protein environment presented as point charges resulted in minimal difference in the vertical excitation energy of the S_1 state by 0.09 eV compared to the isolated

chromophore. Additionally, comparing excited-state descriptors calculated for the chromophore with and without inclusion of the environment revealed that the effect of the environment is minimal. The electron−hole correlation plots for various fragmentation schemes supported by the SA-NTOs show that the first two bright excited states are delocalized, across the whole chromophore, excitations. Calculated excited-state descriptors across the MD trajectory are in line with the results obtained for the representative structure of the MD run. Our findings contribute to the fundamental understanding of the excited-state processes of iRFPs and pave the way toward a molecular-level understanding of the excited-state processes, e.g., ESPT and photoisomerization, which will allow researchers to design new iRFP-based probes with desired spectral properties used as optogenetic tools. Exploring excited-state processes of iRFPs is currently in progress in our group.

■ ASSOCIATED CONTENT

Ⓢⓘ Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpcb.0c10466.

> BV forcefield parameters; hydrogen-bonding pairs of BV−protein in the crystal; clustering analysis of the chromophore; excited-state analysis of BV-I and BV-II models using the CAMB3LYP/cc-pVDZ level of theory; electron−hole correlation plots of BV-I and BV-II models computed at the ωB97X-D/cc-pVDZ level of theory; excited-state analysis of other two predominant clusters; and Ω-matrix plots (PDF)

■ AUTHOR INFORMATION

Corresponding Author

> **Shirin Faraji** − *Theoretical Chemistry Group, Zernike Institute for Advanced Materials, University of Groningen, Groningen 9747AG, The Netherlands*; Email: s.s.faraji@rug.nl

Authors

> **Sivasudhan Rathnachalam** − *Theoretical Chemistry Group, Zernike Institute for Advanced Materials, University of Groningen, Groningen 9747AG, The Netherlands*; ⓘ orcid.org/0000-0002-0016-1387
> **Maximilian F. S. J. Menger** − *Theoretical Chemistry Group, Zernike Institute for Advanced Materials, University of Groningen, Groningen 9747AG, The Netherlands*; ⓘ orcid.org/0000-0003-1442-9601

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.jpcb.0c10466

Notes

The authors declare no competing financial interest.

■ REFERENCES

(1) Rodriguez, E. A.; Campbell, R. E.; Lin, J. Y.; Lin, M. Z.; Miyawaki, A.; Palmer, A. E.; Shu, X.; Zhang, J.; Tsien, R. Y. The Growing and Glowing Toolbox of Fluorescent and Photoactive Proteins. *Trends Biochem. Sci.* **2017**, *42*, 111−129.

(2) Day, R. N.; Schaufele, F. Fluorescent protein tools for studying protein dynamics in living cells: a review. *J. Biomed. Opt.* **2008**, *13*, No. 031202.

(3) Chalfie, M.; Tu, Y.; Euskirchen, G.; Ward, W. W.; Prasher, D. C. Green fluorescent protein as a marker for gene expression. *Science* **1994**, *263*, 802−805.

(4) Tsien, R. Y. The green fluorescent protein. *Annu. Rev. Biochem.* **1998**, *67*, 509−544.

(5) Remington, S. J. Green fluorescent protein: a perspective. *Protein Sci.* **2011**, *20*, 1509−1519.

(6) Bravaya, K. B.; Grigorenko, B. L.; Nemukhin, A. V.; Krylov, A. I. Quantum chemistry behind bioimaging: insights from ab initio studies of fluorescent proteins and their chromophores. *Acc. Chem. Res.* **2012**, *45*, 265−275.

(7) Olenych, S. G.; Claxton, N. S.; Ottenberg, G. K.; Davidson, M. W. The fluorescent protein color palette. *Curr. Protoc. Cell Biol.* **2006**, *33*, 21−25.

(8) Day, R. N.; Davidson, M. W. The fluorescent protein palette: tools for cellular imaging. *Chem. Soc. Rev.* **2009**, *38*, 2887−2921.

(9) Sample, V.; Newman, R. H.; Zhang, J. The structure and function of fluorescent proteins. *Chem. Soc. Rev.* **2009**, *38*, 2852−2864.

(10) Shcherbo, D.; et al. Bright far-red fluorescent protein for whole-body imaging. *Nat. Methods* **2007**, *4*, 741−746.

(11) Helmchen, F.; Denk, W. Deep tissue two-photon microscopy. *Nat. Methods* **2005**, *2*, 932−940.

(12) Baloban, M.; Shcherbakova, D. M.; Pletnev, S.; Pletnev, V. Z.; Lagarias, J. C.; Verkhusha, V. V. Designing brighter near-infrared fluorescent proteins: insights from structural and biochemical studies. *Chem. Sci.* **2017**, *8*, 4546−4557.

(13) Chernov, K. G.; Redchuk, T. A.; Omelina, E. S.; Verkhusha, V. V. Near-infrared fluorescent proteins, biosensors, and optogenetic tools engineered from phytochromes. *Chem. Rev.* **2017**, *117*, 6423−6446.

(14) Li, L.; Shemetov, A. A.; Hu, P.; Shcherbakova, D. M.; Shi, J.; Yao, J.; Verkhusha, V. V.; Wang, L. V. In vivo photoacoustic multi-contrast imaging and detection of protein interactions using a small near-infrared photochromic protein. *Photons Plus Ultrasound: Imaging and Sensing*; International Society For Optics And Photonics 2019. *2019*; pp. 76−81.

(15) Matlashov, M. E.; Shcherbakova, D. M.; Alvelid, J.; Baloban, M.; Pennacchietti, F.; Shemetov, A. A.; Testa, I.; Verkhusha, V. V. A set of monomeric near-infrared fluorescent proteins for multicolor imaging across scales. *Nat. Commun.* **2020**, *11*, 239.

(16) Toh, K. C.; Stojković, E. A.; van Stokkum, I. H. M.; Moffat, K.; Kennis, J. T. M. Proton-transfer and hydrogen-bond interactions determine fluorescence quantum yield and photochemical efficiency of bacteriophytochrome. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 9170−9175.

(17) Shcherbakova, D. M.; Shemetov, A. A.; Kaberniuk, A. A.; Verkhusha, V. V. Natural photoreceptors as a source of fluorescent proteins, biosensors, and optogenetic tools. *Annu. Rev. Biochem.* **2015**, *84*, 519−550.

(18) Shcherbakova, D. M.; Baloban, M.; Emelyanov, A. V.; Brenowitz, M.; Guo, P.; Verkhusha, V. V. Bright monomeric near-infrared fluorescent proteins as tags and biosensors for multiscale imaging. *Nat. Commun.* **2016**, *7*, 12405.

(19) Takala, H.; Björling, A.; Berntsson, O.; Lehtivuori, H.; Niebling, S.; Hoernke, M.; Kosheleva, I.; Henning, R.; Menzel, A.; Ihalainen, J. A.; et al. Signal amplification and transduction in phytochrome photosensors. *Nature* **2014**, *509*, 245−248.

(20) Burgie, E. S.; Wang, T.; Bussell, A. N.; Walker, J. M.; Li, H.; Vierstra, R. D. Crystallographic and electron microscopic analyses of a bacterial phytochrome reveal local and global rearrangements during photoconversion. *J. Biol. Chem.* **2014**, *289*, 24573−24587.

(21) González, R.; Mroginski, M. A. Fully Quantum Chemical Treatment of Chromophore−Protein Interactions in Phytochromes. *J. Phys. Chem. B* **2019**, *123*, 9819−9830.

(22) Feliks, M.; Lafaye, C.; Shu, X.; Royant, A.; Field, M. Structural Determinants of Improved Fluorescence in a Family of Bacteriophytochrome-Based Infrared Fluorescent Proteins: Insights from Continuum Electrostatic Calculations and Molecular Dynamics Simulations. *Biochemistry* **2016**, *55*, 4263−4274.

(23) Modi, V.; Donnini, S.; Groenhof, G.; Morozov, D. Protonation of the Biliverdin IXα Chromophore in the Red and Far-Red Photoactive States of a Bacteriophytochrome. *J. Phys. Chem. B* **2019**, *123*, 2325−2334.

(24) Falklöf, O.; Durbeej, B. Computational Identification of Pyrrole Ring C as the Preferred Donor for Excited-State Proton Transfer in Bacteriophytochromes. *ChemPhotoChem* **2018**, *2*, 453−457.

(25) Zhu, J.; Shcherbakova, D. M.; Hontani, Y.; Verkhusha, V. V.; Kennis, J. T. M. Ultrafast excited-state dynamics and fluorescence deactivation of near-infrared fluorescent proteins engineered from bacteriophytochromes. *Sci. Rep.* **2015**, *5*, 12840.

(26) Hontani, Y.; Shcherbakova, D. M.; Baloban, M.; Zhu, J.; Verkhusha, V. V.; Kennis, J. T. M. Bright blue-shifted fluorescent proteins with Cys in the GAF domain engineered from bacterial phytochromes: Fluorescence mechanisms and excited-state dynamics. *Sci. Rep.* **2016**, *6*, 37362.

(27) Toh, K. C.; Stojković, E. A.; van Stokkum, I. H. M.; Moffat, K.; Kennis, J. T. M. Fluorescence quantum yield and photochemistry of bacteriophytochrome constructs. *Phys. Chem. Chem. Phys.* **2011**, *13*, 11985−11997.

(28) Khrenova, M. G.; Kulakova, A. M.; Nemukhin, A. V. Competition between two cysteines in covalent binding of biliverdin to phytochrome domains. *Org. Biomol. Chem.* **2018**, *16*, 7518−7529.

(29) Polyakov, I. V.; Grigorenko, B. L.; Mironov, V. A.; Nemukhin, A. V. Modeling structure and excitation of biliverdin-binding domains in infrared fluorescent proteins. *Chem. Phys. Lett.* **2018**, *710*, 59−63.

(30) Meech, S. R. Excited state reactions in fluorescent proteins. *Chem. Soc. Rev.* **2009**, *38*, 2922−2934.

(31) Nemukhin, A. V.; Grigorenko, B. L.; Khrenova, M. G.; Krylov, A. I. Computational Challenges in Modeling of Representative Bioimaging Proteins: GFP-Like Proteins, Flavoproteins, and Phytochromes. *J. Phys. Chem. B* **2019**, *123*, 6133−6149.

(32) Topol, I.; Collins, J.; Savitsky, A.; Nemukhin, A. Computational strategy for tuning spectral properties of red fluorescent proteins. *Biophys. Chem.* **2011**, *158*, 91−95.

(33) Faraji, S.; Krylov, A. I. On the Nature of an Extended Stokes Shift in the mPlum Fluorescent Protein. *J. Phys. Chem. B* **2015**, *119*, 13052−13062.

(34) Slavov, C.; Fischer, T.; Barnoy, A.; Shin, H.; Rao, A. G.; Wiebeler, C.; Zeng, X.; Sun, Y.; Xu, Q.; Gutt, A.; et al. The interplay between chromophore and protein determines the extended excited state dynamics in a single-domain phytochrome. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117*, 16356−16362.

(35) Gozem, S.; Luk, H. L.; Schapiro, I.; Olivucci, M. Theory and simulation of the ultrafast double-bond isomerization of biological chromophores. *Chem. Rev.* **2017**, *117*, 13502−13565.

(36) Macaluso, V.; Cupellini, L.; Salvadori, G.; Lipparini, F.; Mennucci, B. Elucidating the role of structural fluctuations, and intermolecular and vibronic interactions in the spectroscopic response of a bacteriophytochrome. *Phys. Chem. Chem. Phys.* **2020**, *22*, 8585−8594.

(37) Plasser, F.; Wormit, M.; Dreuw, A. New tools for the systematic analysis and visualization of electronic excitations I. Formalism. *J. Chem. Phys.* **2014**, *141*, No. 024106.

(38) Plasser, F.; Bäppler, S. A.; Wormit, M.; Dreuw, A. New tools for the systematic analysis and visualization of electronic excitations. II. Applications. *J. Chem. Phys.* **2014**, *141*, No. 024107.

(39) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. Gromacs: High performance molecular

simulations through multi-level parallelism from laptops to super-computers. *SoftwareX* **2015**, *1-2*, 19−25.

(40) 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.

(41) Rao, A. G.; Wiebeler, C.; Sen, S.; Cerutti, D. S.; Schapiro, I. Histidine Protonation Controls Structural Heterogeneity in the Cyanobacteriochrome AnPixJg2. *bioRxiv*, **2020**.

(42) Mark, P.; Nilsson, L. Structure and Dynamics of the TIP3P, SPC, and SPC/E Water Models at 298 K. *J. Phys. Chem. A* **2001**, *105*, 9954−9960.

(43) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: a linear constraint solver for molecular simulations. *J. Comput. Chem.* **1997**, *18*, 1463−1472.

(44) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A smooth particle mesh Ewald method. *J. Chem. Phys.* **1995**, *103*, 8577−8593.

(45) Shirts, M. R.; Mobley, D. L.; Chodera, J. D.; Pande, V. S. Accurate and efficient corrections for missing dispersion interactions in molecular simulations. *J. Phys. Chem. B* **2007**, *111*, 13052−13063.

(46) Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **2007**, *126*, No. 014101.

(47) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **1984**, *81*, 3684−3690.

(48) Nosé, S.; Klein, M. L. Constant pressure molecular dynamics for molecular systems. *Mol. Phys.* **2006**, *50*, 1055−1076.

(49) Parrinello, M.; Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **1981**, *52*, 7182−7190.

(50) Daura, X.; Gademann, K.; Jaun, B.; Seebach, D.; Van Gunsteren, W. F.; Mark, A. E. Peptide folding: when simulation meets experiment. *Angew. Chem., Int. Ed.* **1999**, *38*, 236−240.

(51) Grimme, S. Density functional theory with London dispersion corrections. *WIREs Comput. Mol. Sci.* **2011**, *1*, 211−228.

(52) Mewes, S. A.; Plasser, F.; Dreuw, A. Communication: Exciton analysis in time-dependent density functional theory: How functionals shape excited-state characters. *J. Chem. Phys.* **2015**, *143*, 171101.

(53) Mewes, S. A.; Plasser, F.; Krylov, A.; Dreuw, A. Benchmarking Excited-State Calculations Using Exciton Properties. *J. Chem. Theory Comput.* **2018**, *14*, 710−725.

(54) Mewes, S. A.; Dreuw, A. Density-based descriptors and exciton analyses for visualizing and understanding the electronic structure of excited states. *Phys. Chem. Chem. Phys.* **2019**, *21*, 2843−2856.

(55) Senn, H. M.; Thiel, W. QM/MM Methods for Biomolecular Systems. *Angew. Chem., Int. Ed.* **2009**, *48*, 1198−1229.

(56) Shao, Y.; et al. Advances in molecular quantum chemistry contained in the Q-Chem 4 program package. *Mol. Phys.* **2015**, *113*, 184−215.

(57) Plasser, F. TheoDORE: A toolbox for a detailed and automated analysis of electronic excited state computations. *J. Chem. Phys.* **2020**, *152*, No. 084108.

(58) Humphrey, W.; Dalke, A.; Schulten, K. VMD − Visual Molecular Dynamics. *J. Mol. Graph.* **1996**, *14*, 33−38.

(59) Mewes, S. A.; Plasser, F.; Dreuw, A. Universal exciton size in organic polymers is determined by nonlocal orbital exchange in time-dependent density functional theory. *J. Phys. Chem. Lett.* **2017**, *8*, 1205−1210.

■ NOTE ADDED AFTER ASAP PUBLICATION

Published ASAP on February 24, 2021; Revised March 11, 2021 with citation to ref 41.