JCTC
Journal of Chemical Theory and Computation

Article

What is the Optimal Size of the Quantum Region in Embedding Calculations of Two-Photon Absorption Spectra of Fluorescent Proteins?

Dawid Grabarek and Tadeusz Andruniów*

Cite This: *J. Chem. Theory Comput.* 2020, 16, 6439−6455

Read Online

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

ABSTRACT: We systematically investigate an impact of the size and content of a quantum (QM) region, treated at the density functional theory level, in embedding calculations on one- (OPA) and two-photon absorption (TPA) spectra of the following fluorescent proteins (FPs) models: *Aequorea victoria* green FP (avGFP) with neutral (avGFP-n) and anionic (avGFP-a) chromophore as well as Citrine FP. We find that amino acid (a.a.) residues as well as water molecules hydrogen-bonded (h-bonded) to the chromophore usually boost both OPA and TPA processes intensity. The presence of hydrophobic a.a. residues in the quantum region also non-negligibly affects both absorption spectra but decreases absorption intensity. We conclude that to reach a quantitative description of OPA and TPA spectra in multiscale modeling of FPs, the quantum region should consist of a chromophore and most of a.a. residues and water molecules in a radius of 0.30−0.35 nm (*ca.* 200−230 atoms) when the remaining part of the system is approximated by the electrostatic point-charges. The optimal size of the QM region can be reduced to 80−100 atoms by utilizing a more advanced polarizable embedding model. We also find components of the QM region that are specific to a FP under study. We propose that the F165 a.a. residue is important in tuning the TPA spectrum of avGFP-n but not other investigated FPs. In the case of Citrine, Y203 and M69 a.a. residues must definitely be part of the QM subsystem. Furthermore, we find that long-range electrostatic interactions between the QM region and the rest of the protein cannot be neglected even for the most extensive QM regions (*ca.* 350 atoms).
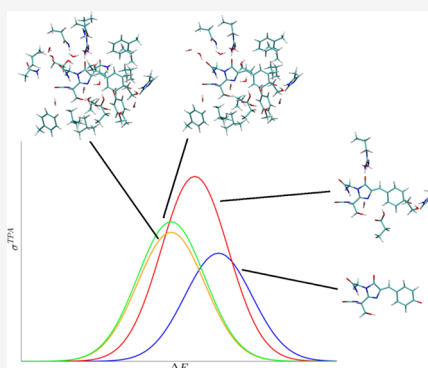
1. INTRODUCTION

Since early works[1,2] on the *Aequorea victoria* jellyfish bioluminescence system, fluorescent proteins (FPs) became a versatile tool in modern biology and biochemistry. They are utilized as biosensors and for the visualization of various processes *in vivo*.[3−5] The vital part of each FP is a chromophore created autocatalytically from three consecutive amino acids and embedded in a characteristic β-barrel polypeptide fold consisting of 11 β-strands (Scheme 1). Currently available FPs differ in terms of spectral, photochemical, and photophysical properties which are dictated by: (i) chromophore structure and (ii) chromophore's protein environment. The latter means that introducing mutation(s) in the a.a. sequence may lead to significant changes in chromophore−environment interactions and hence quantitative and/or qualitative modification of absorption, excitation, and fluorescence spectra which are well covered in various review articles.[3−5] By means of genetic engineering, a colourful palette of FPs was obtained with their absorption and emission spectra spanning the whole visible spectrum as well as near ultraviolet and near infrared.

Recently, FPs gain more attention in techniques based on two-photon absorption (TPA) process, for example, two-photon laser scanning microscopy,[6,7] photodynamic therapy,[5]

Scheme 1. Tertiary avGFP Structure (PDB Code: 1GFL) with the Chromophore in Neutral Protonation State Emphasized
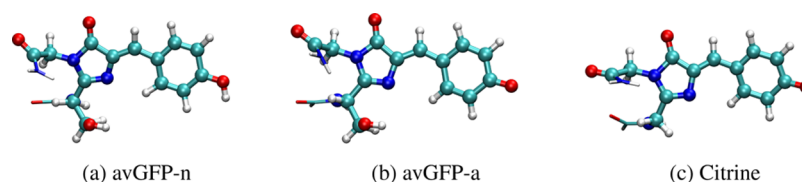
| (a) avGFP-n | (b) avGFP-a | (c) Citrine |
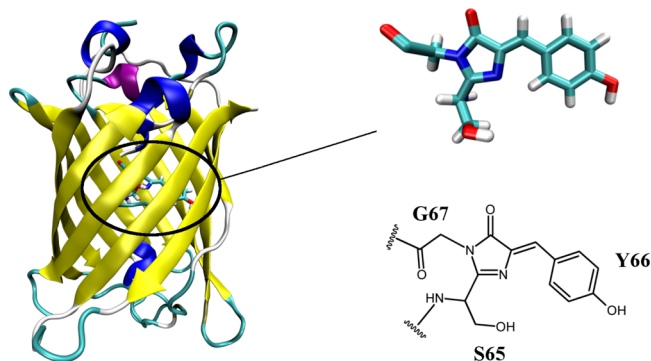
Figure 1. Structure of chromophores in investigated FPs. The distance between atoms in ball and stick representation and the chromophore's environment atoms are used to create the *X.XX* family of QM clusters—see main text for more details. The hydrogen link atoms and carbonyl/amide group of preceeding/following a.a. residues are shown in licorice.

or three-dimensional optical memories.[8] FPs optimized for application in TPA-based techniques should exhibit profound TPA cross-section (σ^{TPA}) value. The consistent experimental TPA spectra measurements revealed a great variation of the σ^{TPA} value between FPs even those possessing the same chromophore structure but different a.a. sequences.[9] For instance, for four representatives of mFruits FP family, the σ^{TPA} value is in the range of 20−44 GM.[9] This clearly shows that one may engineer novel FPs with enhanced TPA cross section. The rational and directed design of FPs requires an understanding of the chromophore—environment coupling impact on its σ^{TPA} value. Recently, Drobizhev and co-workers[10] developed a model that allows to correlate σ^{TPA} with excitation energy (ΔE) for a series of GFP mutants. However, considering their assumptions, it is useful only for the $S_0 \rightarrow S_1$ transition which exhibits the strongest OPA intensity in FPs. However, it was predicted theoretically[11] and later on confirmed experimentally[9,12] that many FPs exhibit much stronger TPA for the $S_0 \rightarrow S_n$ transitions where S_n denotes excited states higher in energy than S_1. Besides, TPA cross-section measurements accuracy suffers from accompanying nonlinear processes,[13] photobleaching,[14] or uncertain mature FP concentration. They may artificially over- or underestimate the true TPA cross section. As a consequence, the reported σ^{TPA} values for enhanced green FP (EGFP) differ even by two orders of magnitude: 1.5,[15] 20,[16] 39,[9] 40,[17] 60,[18] 180,[6,19] and 300 GM[20] for mEGFP (monomeric EGFP) carrying mutations that affect only protein's oligomerization state.

Molecular modeling methods may facilitate the directed development of FPs optimized for TPA-based applications. According to theoretical studies,[21,22] the σ^{TPA} value for the $S_0 \rightarrow S_1$ transition strongly depends on the direction and magnitude of the electric field in the cavity where the chromophore resides which is consistent with results from all-optical experiments.[10,23] More precisely, it is predicted that the σ^{TPA} value can be enhanced by a larger permanent dipole moment change ($\Delta\mu$) upon excitation, that is, more extensive charge transfer. Although tempting and important, this result relies on a two-state model of the TPA process. According to our[24] and other[25] calculations on various FP chromophores in the gas phase, this model is applicable only to the TPA process to the S_1 state. In case of higher ones,[24] a complex channel interference phenomenon[26−28] prohibits a simple correlation between σ^{TPA} and one well-defined electronic feature such as $\Delta\mu$.

Because the $S_0 \rightarrow S_n$ transitions may benefit from large σ^{TPA} values, it is important to characterize chromophore's protein environment features that may increase the TPA cross section. In order to achieve this goal, a reliable protein model is required to be utilized in hybrid QM/MM[29] calculations. In case the of FPs, a chromophore and eventually its immediate environmental make up the QM subsystem which is embedded in the MM (molecular mechanics) environment. In the simplest version of an electrostatic embedding (EE) scheme, electric point-charges are assigned in the position of MM atoms accounting for the QM subsystem electron density polarization because of the electrostatic interaction with the rest of protein. The EE scheme can be extended by adding higher localized multipoles (dipoles, quadrupoles, etc.) along with point-charges. In a more elaborate polarizable embedding (PE) scheme, the QM and MM subsystems polarize each other through placing atom-centered multipoles and polarizabilities in the position of MM subsystem atoms.

The FP model in QM/MM calculations must be carefully chosen to obtain reliable results at the lowest computational cost possible. Steindal et al.[22] calculated excitation energy, TPA cross section, and OPA oscillator strength (f) for avGFP with the chromophore in neutral (avGFP-n) and anionic (avGFP-a) protonation states (Figure 1) using different levels of approximation to the protein model. They conclude that higher order terms in the multipole expansion in EE have a rather small impact on all calculated spectral features for the neutral chromophore case while for the anionic one they are more significant. Introduction of polarization between QM and MM subsystems through anisotropic polarizabilities affects moderately excitation energy and sizably OPA and TPA intensities.[22] Isotropic polarizabilites capture most of the effect though their impact is quantitatively slightly smaller. On the other hand, in some FPs, a sizable impact of polarization interactions on ΔE may be detected.[30] This is particularly the case when coupling between the MM region polarization and electronic excitation is accounted for, using an electronic-state-specific response of the environment to the excitation in the QM region.[30,31]

All these results clearly show that polarization interactions cannot be neglected in calculations of OPA and TPA spectra as also shown for different avGFP models in other FPs.[21,31−33] In fact, the excitation energy for avGFP-a is converged (does not change significantly) only if polarization interactions with a.a. residues and water molecules within at least 20 Å radius from the chromophore are included.[33,34] Surprisingly the convergence radius is somehow smaller for f (18 Å) and σ^{TPA} (14 Å) quantities.[34]

The size of the QM subsystem is also of great importance in OPA and TPA spectra calculations. Kongsted and co-workers[33] have shown that ΔE is considerably red-shifted, almost 0.3 eV, when the QM subsystem is expanded starting from the chromophore alone up to the chromophore + nine nearby a.a. residues and four water molecules. The redshift is visibly smaller for avGFP-a (below 0.1 eV). The largest change in ΔE for both the chromophore protonation states is observed if the three water molecules directly h-bonded to the chromophore become part of the QM subsystem.[33] Furthermore, other authors report that expanding the QM
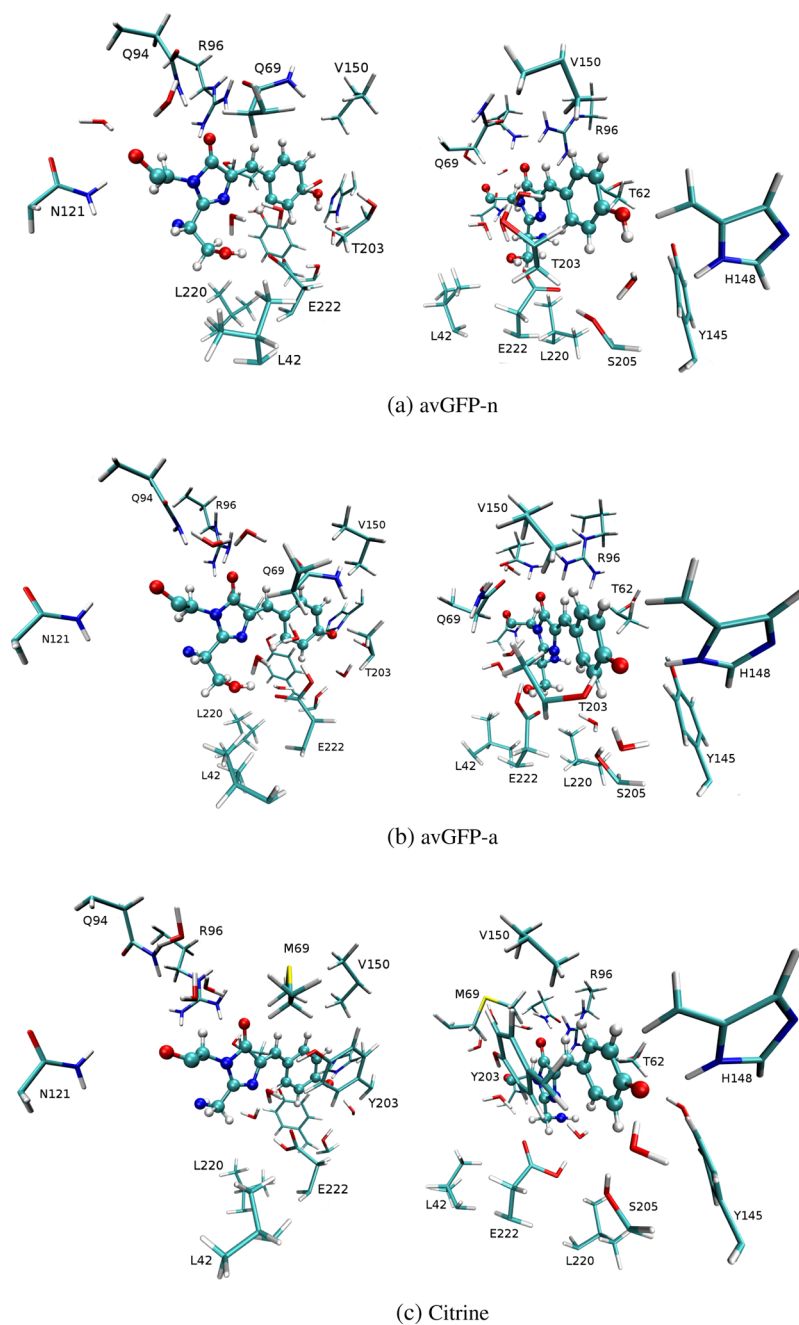
(a) avGFP-n

(b) avGFP-a

(c) Citrine

Figure 2. Chromophore (emphasized in ball and stick representation) and its immediate environment (chosen a.a. residues side chain and water molecules). For each protein, snapshots from two view points are given.

region by seven crystallographic water molecules, residing near the neutral chromophore, leads to a tremendous increase in σ^{TPA} from 16.8 to 52.7 GM and f from 0.444 to 0.711.[22] This clearly shows that either water molecules are involved in charge-transfer upon excitation (which simply cannot be described if they are part of the MM subsystem) or that the PE does not correctly describe the electric field from water molecules interacting with the chromophore. For avGFP-a, the level of theory used to describe water molecules has somehow a more modest impact on OPA and TPA intensities. A systematic study on the impact of the QM subsystem size in avGFP, hosting an anionic chromophore, on its OPA properties revealed that a spectrum is converged when the QM subsystem is made of about 300 atoms, if the rest of the

protein is approximated by EE.[35] This constitutes the chromophore, all a.a. residues and water molecules h-bonded to the chromophore and a few nearby hydrophobic residues. Interestingly, for small QM subsystems (up to *ca.* 110 atoms), including EE leads to a blueshift up to 0.1 eV while for larger QM, the electrostatic field from the rest of the protein has a modest impact on the OPA spectrum. This clearly suggests that h-bonding interactions between the chromophore and its environment are the most relevant ones for tuning ΔE and f at least for that system.[35] However, utilizing PE for the MM subsystem allows to use about 3 times smaller QM subsystem to obtain ΔE and f values that do not change significantly when the QM subsystem size is increased further. This can be

understood as PE being a good approximation of most of the a.a. residues as found previously by Steindal et al.[22]

It seems that choice of the FP model for OPA spectrum calculations is well investigated. This is, however, not the case for the σ^{TPA} value. List et al.[21] compared TPA cross-section values for DsRed FP obtained for the minimal QM subsystem (chromophore only; 63 atoms) with an extended one (242 atoms). They observe that upon extending the QM subsystem size, the σ^{TPA} drops from 105.9 to 74.6 GM. This is a significant difference but the only conclusion is that the QM subsystem size is decisive for results of TPA cross-section calculations. Using an optimal QM subsystem size is important for a good balance between accuracy and required computational resources (memory, cpu or gpu time, number of processors). This is especially the case for σ^{TPA} calculations based on quadratic response theory which are much more expensive than ΔE or f calculations, based on linear response theory. For these reasons, in the present contribution we focus on investigating the relationship between the σ^{TPA} value and the QM subsystem size for the $S_0 \rightarrow S_1$ and $S_0 \rightarrow S_n$ transitions in three representatives of FPs: avGFP with neutral and anionic chromophores[1,2] as well as yellow FP (YFP) Citrine.[36] In avGFP, there is an equilibrium between two protonation states of the chromophore with a 6:1 ratio between neutral and anionic forms near pH ranging from 7 to 8.[37−41] Citrine holds an anionic chromophore similar to that of avGFP, but it is π-stacked with Y203 (numeration of a.a. positions as in avGFP). This $\pi-\pi$ interaction is not present in avGFP and it is mostly responsible for differences in OPA and TPA spectra between the two FPs.[42−44] Furthermore, as we increase the QM subsystem size, the rest of the protein is: (i) neglected, (ii) approximated by EE, or (iii) by PE. We calculate the ΔE, σ^{TPA}, and f quantities describing OPA and TPA spectra using time-dependent density functional theory (TDDFT) with BHandH-LYP[45] functional. According to our benchmark study,[46] this exchange−corelation functional (XCF) above-mentioned spectroscopic quantities in a very good agreement with approximated second-order coupled cluster (CC2) method results. We believe that our results will be a source of firm reference data for choosing the QM subsystem composition for calculations of OPA and TPA spectra in FPs. This will lead to FP models that are reliable for analyzing the impact of a.a. mutations on OPA and TPA spectra which in turn should bring us closer to the rational design of novel FPs with an enhanced σ^{TPA} value.

The remainder of this article is organized as follows: in Section 2, we describe our computational protocol. In Section 3, we present and discuss our results: calculated OPA (ΔE, f) and TPA (σ^{TPA}) properties, in Subsections 3.1 and 3.2, respectively. Section 4 contains conclusions.

2. METHODOLOGY

2.1. Protein Models. To build our FP models, we have used crystallographic structures (PDB code given in parentheses) of avGFP-n (1EMB)[39] and Citrine (1HUY).[36] The crystallographic structure of avGFP-a is not available. Thus, we have utilized a structure obtained for an avGFP-a/S65T mutant (1EMG)[47] and reverted the S65T mutation to obtain a starting structure for an avGFP-a model, which is an approach practiced by others.[48] In the case of avGFP-a and avGFP-n, the crystallographic position of a.a. residues 1, 230−238 was not resolved while for Citrine this is the case for a.a. residues 231−238. The missing atoms in some residues (as described in PDB

files) were added using the Dunbrack's rotamers library[49] as implemented in UCSF Chimera package.[50] The a.a. sequence of avGFP-a and avGFP-n is the same but the important structural differences are related to the E222 protonation state (see next paragraph) and T203 conformation. In avGFP-a, T203 is h-bonded to the chromophore's phenyl oxygen while in avGFP-n its side chain is *ca.* 120° rotated so that threonine's hydroxyl group points away from the chromophore[39] (Figure 2). Citrine holds the following mutations as compared to avGFP: S65G (the first residue in chromogenic tripeptide), V68L, Q69M, S72A, and T203Y, most of them close to the chromophore.

The hydrogen atoms were added using the pdb2gmx tool of Gromacs 2016.3.[51] In the case of GFPs' and YFPs' representatives with an anionic chromophore, the E222 residue is widely acknowledged to be a proton acceptor from the chromophore.[38−40] Hence, the E222 residue of avGFP-a and Citrine is protonated (neutral) and in avGFP-n it carries a negative charge. Other glutamic acid, aspartic acid, lysine, and arginine residues are charged. All histidine residues are neutral and they were assigned the hydrogen atom on either δ (25, 148, 181, 199, 217) or ϵ (77, 81, 139, 169) nitrogen atom based on the visual inspection of the h-bonds network. This choice is mostly consistent with recent subatomic structures from X-ray and neutron crystallography experiments for avGFP mutants.[52,53] The most important concern is about H148 a.a. residue which interacts directly with the chromophore (Figure 2) and seems to be cationic.[52,53] However, this protonation state may be characteristic for the crystal structure only because of the large steric hindrance between hydrogen bonded to ϵ nitrogen and nearby atoms.[53] Because the reasons behind the newly discovered protonation state of H148 are not fully understood, we choose to use "traditional" H148 in our models with hydrogen bonded to the δ nitrogen atom only. This is consistent with earlier experimental structures[39,47] as well as models used in other theoretical works.[30,31,48,54] Each FP model was soaked in a rectangular TIP3P[55] water box so that the minimum distance between any of a protein atom and box edge is 1.2 nm. The Na$^+$ and Cl$^-$ ions were added to neutralize the whole system and obtain a physiological salt concentration of 0.15 mol/dm^3. These proteins' models were utilized for classical and QM/MM molecular dynamics (MD) simulations, as described in the Subsection 2.2. In the latter case, the QM subsystem is composed of the chromophore, V68 or L68, as well as water molecules and hydrophilic a.a. side chains that are either directly h-bonded to the chromophore or are a part of the complex h-bond network in the immediate chromophore vicinity, as shown in Figure S1 in the Supporting Information. In the case of Citrine FP, the Q94 residue is not included in the QM subsystem because in the course of classical MD simulation, it surfs away from the chromophore (see section S2 in the Supporting Information).

To calculate ΔE, σ^{TPA}, and f, the QM and MM subsystems must be defined. First, the chromophore as shown in Figure 1, consists of a chromogenic triplet extended by carbonyl and amide groups of preceeding and following a.a. residues, respectively, to avoid placing the QM/MM boundary on peptide bonds. We increase QM subsystem size by either including whole water molecules or the side chains of a.a. residues, that is, the bonds between C_α and C_β atoms must be cut to separate QM and MM subsystems (neither proline nor glycine was included in any QM cluster). The QM atoms at the QM/MM boundary were saturated with hydrogen atoms

placed along the cut bond in 0.1 nm distance. The a.a. side chain or water molecule is part of the QM subsystem if at least one of its atoms is within given radius from any of the chromogenic triplet's atom (Figure 1). The radius is changed from 0.20 to 0.50 nm with a 0.05 nm interval and the obtained FP models were named accordingly (X.XX). Apart from that, intermediate models (denoted as *intX*, where X is a natural number) with arbitrary chosen QM subsystem composition were created. This serves to fill in gaps between the *X.XX* family of QM clusters and gain a deeper understanding of the relationship between the QM subsystem size/composition and calculated spectral properties. In our QM clusters, we do not include side chains of F64 and L68 residues covalently bonded to the chromophore in all investigated FPs. This is done to avoid the steric hindrance between hydrogen link atoms used to saturate the cut bonds. The rest of the protein is: neglected or approximated by EE or PE force fields, as described in Subsection 2.3. The composition of all QM subsystems utilized in the stage of spectral properties calculations is provided in Tables S1−S6 in the Supporting Information.

The structural features of FP models utilized in spectral property calculations compared to the crystal structure are available in the Supporting Information.

2.2. Molecular Dynamics. The crystallographic structures of our FP models have been relaxed first using a classical MD simulation with Gromacs 2016.3 computational package.[51] The details of the MD protocol are available in the Supporting Information, and here, we provide a brief description. We have used Amber ff99SB*-ildn force field[56−58] with the parameters for chromophores provided by Nifosí and co-workers.[59] First, the steepest descent algorithm was employed to optimize positions of solvent molecules and protein's hydrogen atoms with heavy atoms restrained at their crystallographic positions. The bond lengths between heavy and hydrogen atoms were frozen; hence, we used a 2 fs time step in MD simulations. We gently heated up the system to 300 K using a Berendsen thermostat. Then, we switched to the *NPT* ensemble and used a Berendsen isotropic barostat with a reference pressure equal 1 bar. The restraints on heavy atom positions were gradually released for 100 ps all together. Next, all restraints on atoms positions were removed and 10 ns long production run was performed in the *NPT* ensemble using a Nosé−Hoover thermostat and Parrinello−Rahman barostat.

Then, we arbitrary selected a snapshot from the last stage of classical MD simulation (after 4.5 ns) for structure refinement with Born−Oppenheimer QM/MM MD in CP2K 2.6.1 package.[60] Hybrid MD simulation was performed in the *NVT* ensemble with time step reduced to 1 fs even though bond lengths between heavy and hydrogen atoms remain frozen. We first heated up the system for 900 fs to 300 K and then performed a production run for 10 ps. The QM subsystem was treated at the DFT/BLYP[61,62] level of theory with the 6-31G* basis set. The electrostatic coupling between QM and MM subsystems was utilized and the Amber force field point charges near the QM/MM boundary were redistributed to avoid overpolarization of electron density in the QM subsystem. Other important details of our QM/MM MD protocol are available in the Supporting Information.

2.3. OPA and TPA Spectrum Calculations. We have calculated spectral properties using Turbomole 7.3[63] [no embedding (NE) and EE] and Dalton 2018.0[64,65] (PE). One-photon quantities (ΔE and f) were calculated using linear response theory[66] and f is given in length representation. The

TPA cross section was calculated within quadratic response theory ansatz.[66,67] We performed all calculations using the TDDFT method with the BHandHLYP[45] hybrid XCF. In calculations with NE and EE, we applied the aug-cc-pVDZ basis set[68] but in PE the cc-pVDZ basis set. Using diffuse basis functions in combination with PE led to various artifacts and we decided to discuss this issue in a paper to come. Nevertheless, the OPA and TPA spectra obtained with and without diffuse basis functions are very similar for excitation energies up to 5.0 eV—the spectral region that includes bright OPA and TPA bands. Hence, in our view, the conclusions regarding the relationship between absorption spectra and QM subsystem size should be the same with aug-cc-pVDZ basis set and PE applied. According to our recent benchmark of XCF functionals, BHandHLYP provides ΔE, σ^{TPA}, and f values in an excellent agreement with the reference CC2 method results for a set of FP chromophores *in vacuo*.[46] One could argue whether a long-range corrected hybrid CAM-B3LYP XCF[69] is a better choice because of possible charge-transfer character of some excited states we investigate. However, the performance of BHandHLYP and CAM-B3LYP is similar according to our benchmark.[46] Second, CAM-B3LYP became available for TPA transition moments calculations in the most recent Turbomole 7.5[70] that was released after the submission of this manuscript. We shortly discuss that trends in OPA and TPA spectra as a function of the QM region size should be analogous with either the BHandHLYP or CAM-B3LYP XCF (Subsection S6.1 in the Supporting Information). We utilized the Grimme's D3 dispersion correction[71] together with Becke−Johnson damping.[72,73] It affects the results at the stage of solving the Kohn−Sham (KS) equations. The resolution of identity approximation for Coulomb integrals (RI-J)[74,75] with the aug-cc-pVDZ auxiliary basis set was used in calculations with Turbomole. For QM clusters composed of *ca.* 120 atoms or more, we decreased the threshold of neglecting the integrals to 10^{-15} because otherwise KS equations would not converge. The default value is $10^{-8}/3 \cdot N_{bf}$, where N_{bf} denotes the number of basis functions. The final value is on the order of 10^{-12}. In the case of smaller QM subsystems, the spectral properties are not affected by this parameter value choice. The number of excited states for which ΔE, σ^{TPA}, and f properties were calculated is given in Tables S10−S18 in the Supporting Information. Our goal was to reach excited states exhibiting profound TPA cross section where possible.

We have analyzed the σ^{TPA} value which is an experimentally measurable quantity while the original result of quadratic response calculations is rotationally averaged TPA transition moment $\langle \delta^{TPA} \rangle$. We converted $\langle \delta^{TPA} \rangle$ to σ^{TPA} as in our previous works.[24,46] The more detailed discussion is available in Subsection 2.5 of ref 46. In short, the formula for σ^{TPA} is

$$\sigma^{TPA} = \frac{4\pi^2 \alpha a_0^5 \omega^2 \langle \delta^{TPA} \rangle}{c\Gamma} \tag{1}$$

where α is the fine structure constant, a_0 is the Bohr radius, ω is the photon energy, c is the speed of light, and Γ is an empirical damping parameter to describe the spectral broadening of the absorption peak. We define Γ as half-width at half-maximum of the absorption peak and we chose it to be 0.1 eV for all excited states as frequently practiced by others.[11,21,22,76−80]

In the EE, the point-charges from the classical Amber force field (the same as used in MD simulations) were placed in

position of MM subsystem atoms. Because the sum of point-charges of a.a. side chains is not an integer in Amber force fields, we would obtain a noninteger total charge for the EE as the a.a. side chains are added to the QM subsystem. This is nonphysical; hence, we evenly redistributed the spurious electric charge to the protein sites remaining in the MM subsystem (*ca.* 3250–3550). It does not significantly affect the final point-charges (on the order of 10^{-5} to 10^{-4}) values. Subsequently, to prevent the QM subsystem overpolarization, all point-charges within 1.4 Å from the QM cluster atoms were set to zero and redistributed to the 3 nearest sites using the PE library as implemented in Dalton 2018.0 package. That was the only place where Dalton was used in setting up the EE calculations.

The PE force field consists of atom-centered static multipoles up to and including quadrupoles and anisotropic dipole–dipole polarizabilities. They were derived using localized property (LoProp) approach[81] as implemented in OpenMolcas[82] at the B3LYP[83]/ANO-L-VDZP[84–86] level of theory. To obtain these electrostatic properties, the protein part belonging to the MM subsystem was divided into separate a.a. residues using molecular fractionation with conjugate cap (MFCC) scheme[87] and the final PE parameters were obtained using the procedure reported by Søderhjelm and Ryde.[88] The MFCC procedure and PE force field assembly was automated with PyFrame0.2.0 package developed by J. M. H. Olsen.[89] The point-charges within 1.4 Å from any of the QM cluster atoms were redistributed as in the EE case, while higher order multipoles and anisotropic polarizabilities were removed. We perform TDDFT/PE calculations as implemented by Kongsted and co-workers[90] with the effective external field (EEF) correction applied.[91] EEF serves to model the MM subsystem polarization because of an external electromagnetic field, for example, light which triggers OPA or TPA process. Because, only the induced electrostatic field intensity but not frequency is modified, the EEF correction does not influence ΔE value. We shortly discuss the EEF correction impact on OPA and TPA spectrum intensities in Subsection S6.3 in the Supporting Information.

As a final remark, in EE we use every MM subsystem water molecule and ion as a source of the electrostatic field. In the PE force field, we take all water molecules that are up to 3 Å from any protein atom to reduce the computational time. This includes water molecules that are trapped inside the protein fold as well as the layer around the protein. Including water molecules within 5 or 7 Å radius in a polarizable force field (Tables S19–S20 in the Supporting Information) is not expected to affect the conclusions in this work as we discuss in Subsection S6.2 in the Supporting Information.

3. RESULTS AND DISCUSSION

All ΔE, σ^{TPA}, and f values obtained with different QM subsystems are available in Tables: S10–S12 (avGFP-n), S13–S15 (avGFP-a), and S16–S18 (Citrine) in the Supporting Information depending on the level of theory used to describe the MM subsystem: NE, EE, and PE. To simplify the reading, we use a following notation to refer to an investigated FP model: qm/mm, where qm is a QM subsystem as defined in Tables S1–S3 (avGFP-n), S4–S6 (avGFP-a), and S7–S9 (Citrine) in the Supporting Information, and mm subsystem is neglected (NE), approximated by EE or PE.

Figure 2 shows the relaxed structure of the chromophore and its closest environment. We display the chosen a.a.

residues and water molecules in order to make it easier for the readers to navigate through the text.

3.1. OPA Spectra. The characteristic feature of the FPs' OPA spectrum is a very bright peak at ca. 3.0–3.2 eV as seen in Figures 3–5 for chosen QM subsystems (for full set of OPA spectra see Figures S4–S5, S7–S8 and S10–S11 in the Supporting Information). Within both NE and EE levels of theory, this band is usually created by a single $S_0 \rightarrow S_1$
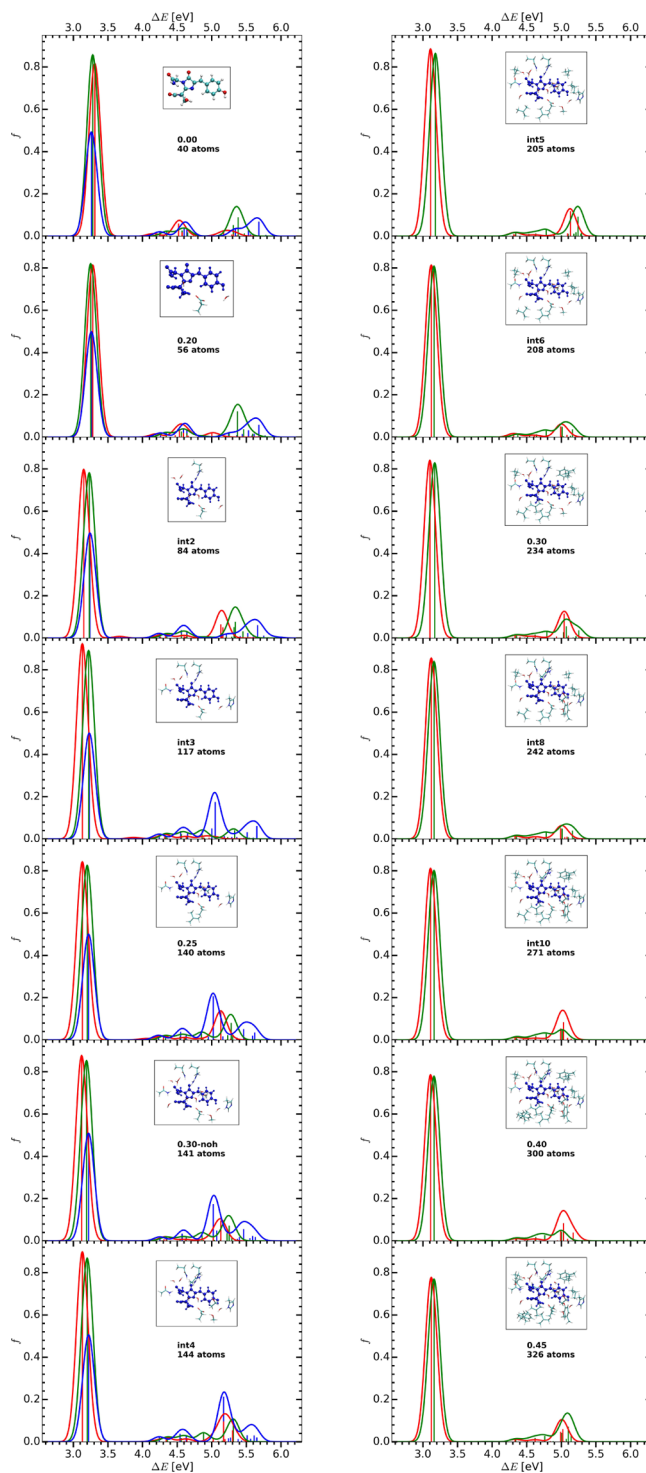

Figure 3. OPA spectra for the chosen QM subsystems of avGFP-n models. The NE, EE, and PE results are given in red, green, and blue, respectively.
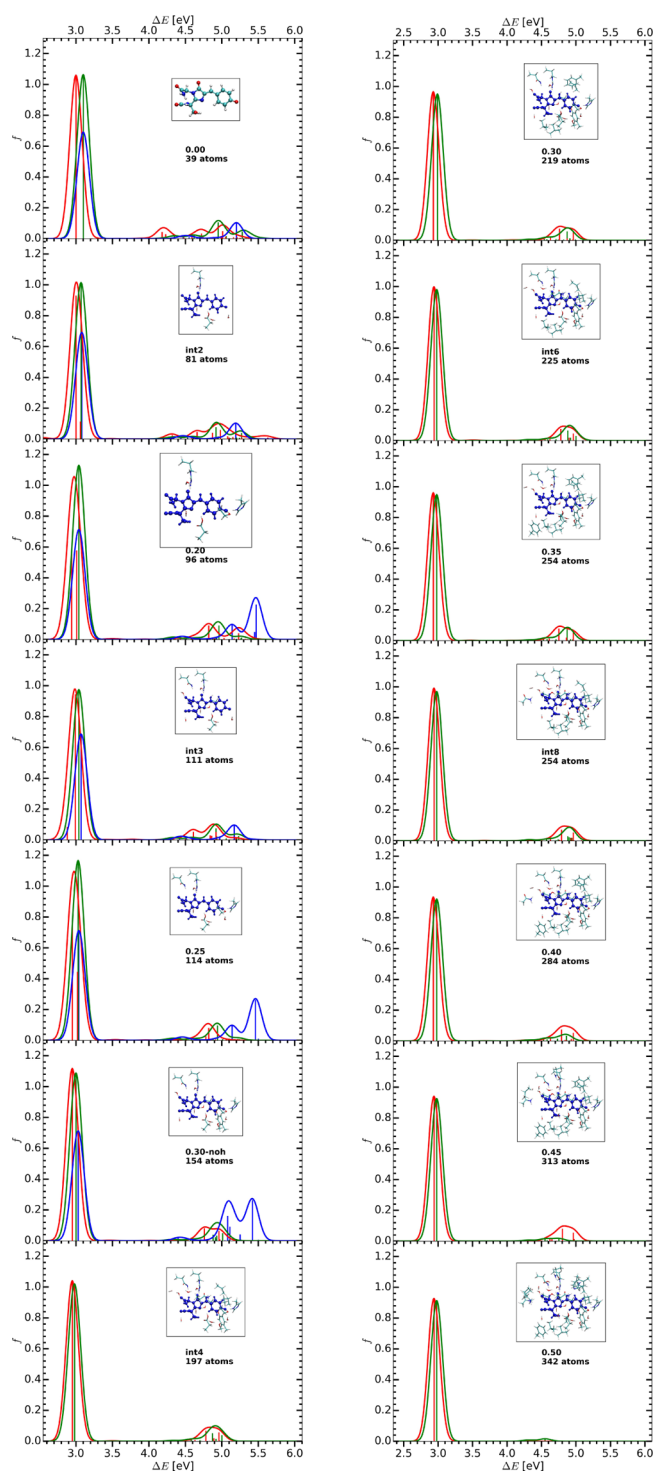
Figure 4. OPA spectra for chosen QM subsystems of avGFP-a models. The NE, EE, and PE results are given in red, green, and blue, respectively.

Figure 5. OPA spectra for chosen QM subsystems of Citrine models. The NE, EE, and PE results are given in red, green, and blue, respectively.

transition of the $\pi \rightarrow \pi^*$ character localized on the chromophore.

3.1.1. avGFP-n. For int1/NE model (0.00/NE + R96 + E222), the above-mentioned excitation pattern does not hold (Figure S4 in the Supporting Information). The brightest OPA peak is created by two excited states that are 0.12−0.19 eV (f equal 0.124 and 0.637, respectively) red-shifted as compared to a single one in 0.00/NE with f of 0.817. Such a split is generally viewed as a result of model shortcomings.[22,35] This is
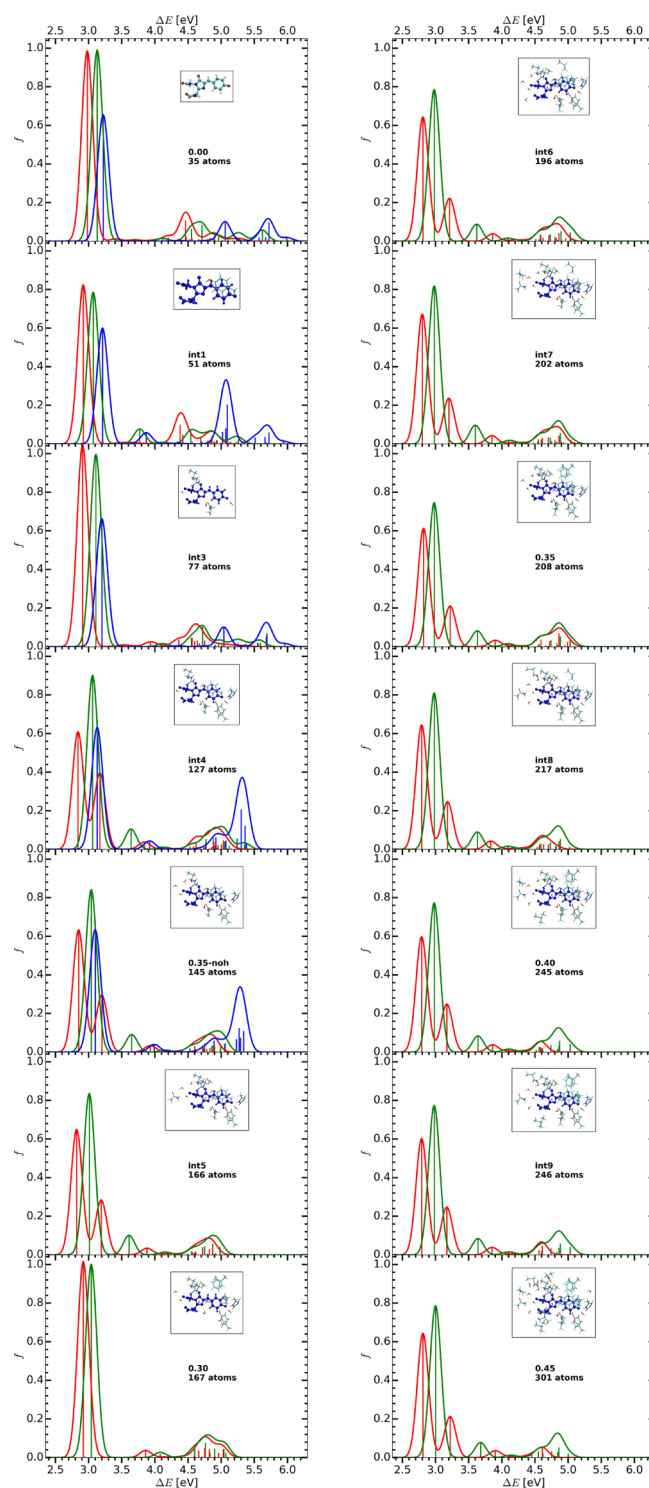
supported by MOs involved in the split transitions being diffused into the environment, that is, not fully localized on the chromophore. The excited state split is suppressed by adding five H_2O molecules to the QM subsystem (int2/NE). Also when the electrostatic interaction between QM and MM regions is accounted for in int1/EE model, we observe one excited state of the $\pi \rightarrow \pi^*$ character. Compared to 0.00/EE, the corresponding excited state is only 0.03 eV red-shifted in

int1/EE and f goes from 0.858 to 0.767. It is noteworthy, that the latter value is almost equal to the sum of OPA oscillator strengths for the two excited states of the int1/NE model.

As the QM size increases, further redshift of the lowest lying OPA peak is observed for both NE and EE cases (Figures 3 and S4 in the Supporting Information). OPA intensity is enhanced by including hydrophilic a.a. residues Q94, N121, and H148 h-bonded to the chromophore (int3; Figure 3). However, f drops down for a larger system (0.25) presumably because of the hydrophobic L220 a.a. residue presence in the QM cluster. For instance, f is larger for 0.30-noh or 0.35-noh (0.30 or 0.35 cluster without hydrophobic a.a. residues) than for 0.30 or 0.35 (Figure S4 in the Supporting Information). Furthermore, in 0.25, 0.30-noh, and int4 series of QM clusters sharing virtually the same size (140−144 atoms) but different compositions, we observe increase in f as only polar a.a. residues are added (ΔE changes by up to 0.01 eV). Starting from int6, as the QM subsystem size increases so does a f value (ΔE changes up to 0.02 eV). This is up to the int8 system (242 atoms) when the f value systematically but slowly decreases with the QM subsystem size. ΔE is converged starting from int6 (but in 0.30 changes by 0.01 eV) if EE is used. By neglecting the electrostatic interactions with the MM subsystem, we cannot say that we reached a converged ΔE value for the S_1 excited state as it oscillates between 3.10 and 3.12 eV for QM systems consisting of at least 205 atoms (Table S10 in the Supporting Information).

The calculated OPA spectrum is much less dependent on the QM cluster size if we apply the polarizable force field (Figures 3 and S5 in the Supporting Information). The oscillator strength for the brightest S_1 excited state is virtually immune to the QM subsystem composition as it changes between 0.492 and 0.509. Similarly, the excitation energy slowly decreases from 3.26 eV (0.00/PE) to 3.22 eV (0.25/PE, 0.30-noh/PE, and int4/PE). Quite contrary, ΔE does not change or blue-shifts by up to 0.03 eV when the QM region of the 0.00/PE model is extended with F64 and V68 a.a. residues covalently bonded to the chromophore as reported by Steindal et al.[92] Also the f value decreases for this larger chromophore model,[92] while we report that adding water molecules and a.a. side chains has rather a reverse impact (Table S12 in the Supporting Information). Apparently, extending the QM region "along covalent bonds" leads to qualitatively different changes in OPA spectrum features. Nevertheless, we would have to perform our calculations for more than one snapshot (Steindal et al. used five and always obtained similar trend) to be more confident about that conclusion. We stress that these spectral features do not change significantly for QM clusters consisting of 140−144 atoms as compared to much smaller int1 (0.00 + R96 + E222), int2 (int1 + 5 × H_2O) ones or even the chromophore alone (Figure 3). This was not the case with EE where the difference between int1/EE and int4/EE models in terms of ΔE and f was 0.05 eV and 0.102, respectively. This error is reduced to merely 0.02 eV and 0.014 with PE applied.

3.1.2. avGFP-a. The int2/NE, 0.20/NE, 0.25/NE, and int3/NE models of avGFP-a are unreliable for OPA calculations because the relevant excited states lose its $\pi \rightarrow \pi^*$ character which is recovered when larger QM regions are in use. This is accompanied by a split of the brightest OPA excited state (Figure S7 and Table S13 in the Supporting Information). It is interesting to note that we find a split for enlarged QM subsystems using NE, while Kongsted and co-workers[35] found a similar split but only for a bare chromophore with EE. This

could be attributed to a difference in the geometry and/or computational details (XCF, basis set). On the contrary, according to our results, using EE or PE never leads to a split and ΔE usually decreases when the QM subsystem size increases for the abovementioned QM subsystems (with an exception of int3/PE) but f displays a more complex behavior (Figures 4 and S7 in the Supporting Information). For instance, for 0.20/EE and 0.25/EE systems f are equal 1.130− 1.167 while for in-between int3/EE system, it is 0.975 (Table S14 in the Supporting Information). This sudden drop in the OPA intensity may be attributed to more water molecules in the int3 system (Table S6 in the Supporting Information). We obtain a similar result with PE although the difference in terms of f is merely 0.025−0.026 (Table S15 in the Supporting Information). In contrast, ΔE is 0.03 eV higher for the int3/PE model as compared to 0.20/PE and 0.25/PE. For these QM clusters and EE model, this difference was up to 0.01 eV. This suggests that water molecules are poorly described by PE.[22] As it was the case for avGFP-n, we observe a red-shift of the brightest OPA peak as the QM region increases while Steindal et al.[92] observed a blue-shift up to 0.06 eV using the 0.00/PE + F64 + V68 model as compared to results for the 0.00/PE structure. Extending the chromophore along peptide bonds has quantitatively similar impact on f value[92] (0.00−0.01) as extending the QM region according to our protocol. Starting from the 0.30-noh system, f displays a decreasing trend with a QM size (Figures 4 and S7 in the Supporting Information) but intX clusters have higher f than X.XX ones of similar size (compare results for int5, int6 with 0.30 and int7, int8 with 0.35 QM clusters in Tables S13 and S14 in the Supporting Information). This is true for both EE and NE cases but cannot be resolved for results obtained with PE as we calculated spectra only for QM subsystems up to 154 atoms. The intX representatives seem to be richer in water molecules and hydrophilic a.a. residues. We observed a similar reason for OPA enhancement in case of avGFP-n.

3.1.3. Citrine. Although Citrine and avGFP-a FPs share a similar chromophore, a T203Y mutation in the former leads to a qualitatively different OPA spectra. This is well illustrated by the int1 system consisting of precisely the chromophore and Y203 (Figures 5 and S10−S11 in the Supporting Information). By applying EE and PE, we find the $S_0 \rightarrow S_1$ transition of the $\pi \rightarrow \pi^*$ character localized on the chromophore as well as the $S_0 \rightarrow S_2$ transition which displays a charge transfer character from Y203 to the chromophore (Tables S17 and S18 in the Supporting Information). The latter has a non-negligible f value of 0.077 and is found 0.70 eV higher in terms of energy than the brightest OPA excited state (S_1) within EE approximation. It is worth to note, that simultaneously the $S_0 \rightarrow S_1$ transition displays a smaller f value (0.786 compared to 0.990 for a bare chromophore). Within the int1/PE model, we observe that the sum of oscillator strengths for the $S_0 \rightarrow S_1$ and $S_0 \rightarrow S_2$ transitions is almost exactly equal to the f value for the $S_0 \rightarrow S_1$ transition as predicted for the 0.00/PE model (Figure 5). If the electrostatic field from the MM subsystem is not included in the int1/NE model, we find the $S_0 \rightarrow S_1$ transition of the $\pi \rightarrow \pi^*$ character and the S_{11} and S_{12} excited states which are dominated by $\pi \rightarrow \pi_{cap}^*$ transition with a small contribution from $\pi_{Y203} \rightarrow \pi^*$ transition (Table S16 in the Supporting Information). π_{cap}^* denotes antibonding π MO localized on the capping moieties of the chromophore (Figure 1). Their f value is 0.099 and 0.047 and ΔE is 4.38 and 4.42 eV, respectively, that is, almost 1.5 eV above the S_1 excited

state. This is *ca.* two times larger difference in terms of ΔE than for the int1/EE case. It takes the int4/NE system consisting of a few a.a. residues h-bonded to the chromophore (Tables S7−S9 in the Supporting Information) to find S_1 and S_3 excited states, both of $\pi + \pi_{Y203} \rightarrow \pi^*$ character, that is, the MO excited from is delocalized over the chromophore and the Y203 a.a. residue (almost perpendicular to the chromophore) and the MO excited into is localized on the chromophore only. When EE or PE is applied, the brightest one-photon transition is still of $\pi + \pi_{Y203} \rightarrow \pi^*$ character but the band at *ca.* 3.6 eV (EE) or *ca.* 3.9−4.0 eV (PE) is created by a $\pi_{Y203} \rightarrow \pi^*$ transition. Interestingly, without any embedding, the ΔE difference between discussed excited states is only *ca.* 0.4 eV but 0.6−0.7 eV using EE or 0.7−0.9 eV using PE (Figure 5). This is because of differential stabilization of discussed excited states by the electric field stemming from the MM region. Also the ratio of f values for the two transitions is visibly different in NE, EE, and PE cases (Figure 5 and Tables S16−S18 in the Supporting Information). Analysis of OPA spectra, within all NE, EE, and PE models (Figures S10−S11 in the Supporting Information), clearly shows that $\pi_{Y203} \rightarrow \pi^*$ transition, responsible for the above-mentioned spectral features, is strongly influenced by interactions with other a.a. residues and water molecules as predicted by Beerepoot et al.[44] We will focus on the results obtained with EE because it leads to a more complete Citrine model than NE and we were able to analyze results for more QM subsystems than with PE. As the QM subsystem size increases starting from int4/EE, both discussed excited states usually become dimmer for the OPA process (Figure 5). We find that int7/EE and int8/EE clusters have a somewhat larger f value than int6/EE and 0.35/EE clusters of a similar size (Figure S10 in the Supporting Information). By analyzing their composition (Tables S7−S9 in the Supporting Information), we could attribute that to the N121 a.a. residue which is present in the OPA brighter clusters. However, if we extend the QM subsystem further to *ca.* 250−350 atoms, we find that f damps and is closer to the one obtained for the int6 cluster composed of only 196 atoms and missing N121. Hence, we do not find this a.a. residue as a key element in shaping OPA in Citrine. Its OPA process enhancement can be attributed to too small QM subsystem in the case of int7/EE and int8/EE models.

As Figures 3−5 reveal, there are also OPA bands in the ultraviolet part of the electromagnetic spectrum (4.5−5.5 eV excitation energies). Combined with significantly lower f value than $S_0 \rightarrow S_1$ transition, the OPA process to these excited states does not seem to be useful in practical applications of FPs. In short, these OPA bands are mostly created by at least two excited states. According to spectra in Figures 3−5, this fragmentation is more visible in the case of EE, at least for of avGFP-n and avGFP-a. It is pretty hard to find any trends in this tangle of OPA bands but we can safely state that we did not manage to reach the converged results with investigated models. In case of the calculations involving PE approximation, we also find a much dimmer OPA process for excited states higher than S_1. However, if the H148 a.a. residue is part of the QM subsystem we observe quite a profound OPA band at ΔE > 5.0 eV (Figures 3−5 and S5, S8, S11 in the Supporting Information). We attribute it to the excitation localized on the H148 a.a. residue itself.

3.2. TPA Spectra. FPs display relatively modest TPA intensity for the S_1 excited state. However, as predicted theoretically[11] and confirmed by the experiment,[9] FPs may benefit from resonant enhancement to produce more intense TPA bands for transitions to higher excited states (S_n). We discuss the σ^{TPA} values for transitions to S_1 and S_n excited states, separately.

3.2.1. $S_0 \rightarrow S_1$ Transition. 3.2.1.1. avGFP-n. For the 0.00/ NE model of avGFP-n, the σ^{TPA} value is modest (2.5 GM) but quickly increases with the system size to reach the values of *ca.* 7−9 GM for QM clusters containing, except chromophore, at least R96, E222, and five water molecules (Figure 6, Table S10
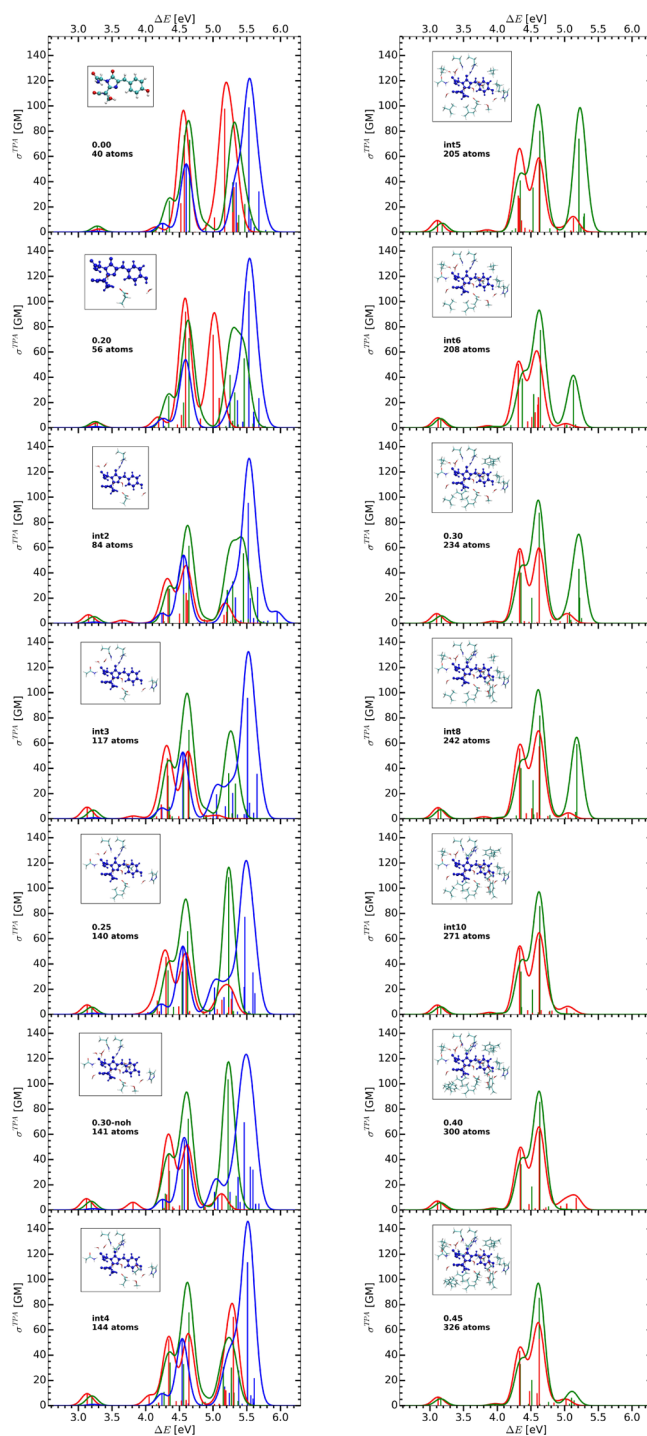

Figure 6. TPA spectra for chosen QM subsystems of avGFP-n models. The NE, EE, and PE results are given in red, green, and blue, respectively.

in the Supporting Information). Starting from the int8/NE model, σ^{TPA} reveals a rather decreasing trend just like f with an increasing QM region (Figure S3 in the Supporting Information). The 0.00/EE model has already two times larger σ^{TPA} than 0.00/NE (Tables S10 and S11 in the Supporting Information). This clearly shows that the presence of an electrostatic field from the remaining part of FP accounts to some extent to the impact of increasing the QM subsystem size on σ^{TPA} in the NE case. As a consequence, extending the QM subsystem with R96, E222, and even five water molecules (0.20/EE−int2/EE) has a minor impact on σ^{TPA}. It takes at least int3/EE model to reach a σ^{TPA} of 7.0 GM which additionally contains N121, Q94, and H148—the last two h-bonded to the chromophore (Figure 2a). This value does not change significantly for the int5/EE−int9/EE family but for 0.25/EE−0.50/EE and int10/EE models, the σ^{TPA} value is again slightly smaller by 1−1.5 GM (Tables S10−S11 in the Supporting Information). The plausible reason for that is the presence of the F165 residue in the QM subsystem in the latter models (excluding 0.25 cluster; Tables S1−S3 in the Supporting Information); however, it is not involved in the $S_0 \rightarrow S_1$ transition in terms of molecular orbitals. A crystal structure, as well as results of our MD simulations, indicates that F165 is not parallel to the chromophore like Y203 in Citrine. We note that for very large QM subsystems, such as 0.45/EE (326 atoms) or 0.50/EE (365 atoms), the σ^{TPA} value is very close to the one for the 0.00/EE model. This is not the case for ΔE and f, however, as we discussed in Section 3.1. On the one hand this may be a coincidence because of error cancellation. It can be also argued that the chromophore—environment interactions are well described by the simplest point-charge model of the environment for σ^{TPA} calculations and it takes an extensive QM subsystem to damp the effect of chromophore—environment interactions when the latter is described at the QM level of theory. This apparently has to do with a major role of electrostatic interactions in tuning TPA spectra in FPs as frequently suggested.[9,10,21,22]

3.2.1.2. avGFP-a. In the case of the anionic chromophore of avGFP-a, we observe that σ^{TPA} reaches the maximum value (5.1−5.7 GM) for 0.25/NE and 0.25/EE systems (Figure 7) containing the chromophore, Q94, R96, H148, T203, E222, and few water molecules in the QM region, which accounts for virtually all species h-bonded to the chromophore (Figure 2b). Then, for 0.30-noh/NE it falls down to the same value as for 0.00/NE (2.4 GM). The 0.30-noh QM cluster contains a complete h-bond network near the chromophore (0.25 + T62 + Y145 + S205 + 3 × H_2O). However, we do not observe significant σ^{TPA} drop when going from 0.25/EE to 0.30-noh/EE models in contrast to 0.25/NE and 0.30-noh/NE ones. Thus, the electrostatic interactions between QM and MM subsystems "stabilize" TPA intensity. With the EE applied, the σ^{TPA} value decreases only slightly (to 3.7−4.8 GM) when the QM subsystem size increases to 197−342 atoms. As it was the case for avGFP-n, the σ^{TPA} value for 0.45/EE−0.50/EE models is close to 0.00/EE, and in the NE case the agreement is even better already for smaller systems such as int4/NE−int9/NE compared to 0.00/NE (Tables S13 and S14 in the Supporting Information).

3.2.1.3. Citrine. For Citrine, we observe that the σ^{TPA} value decreases when Y203 is added to the QM subsystem from 15.5 to 12.4 GM in the NE case (Table S16 in the Supporting Information) and 16.9 to 13.8 GM in the EE case (Table S17 in the Supporting Information). This was similarly true for
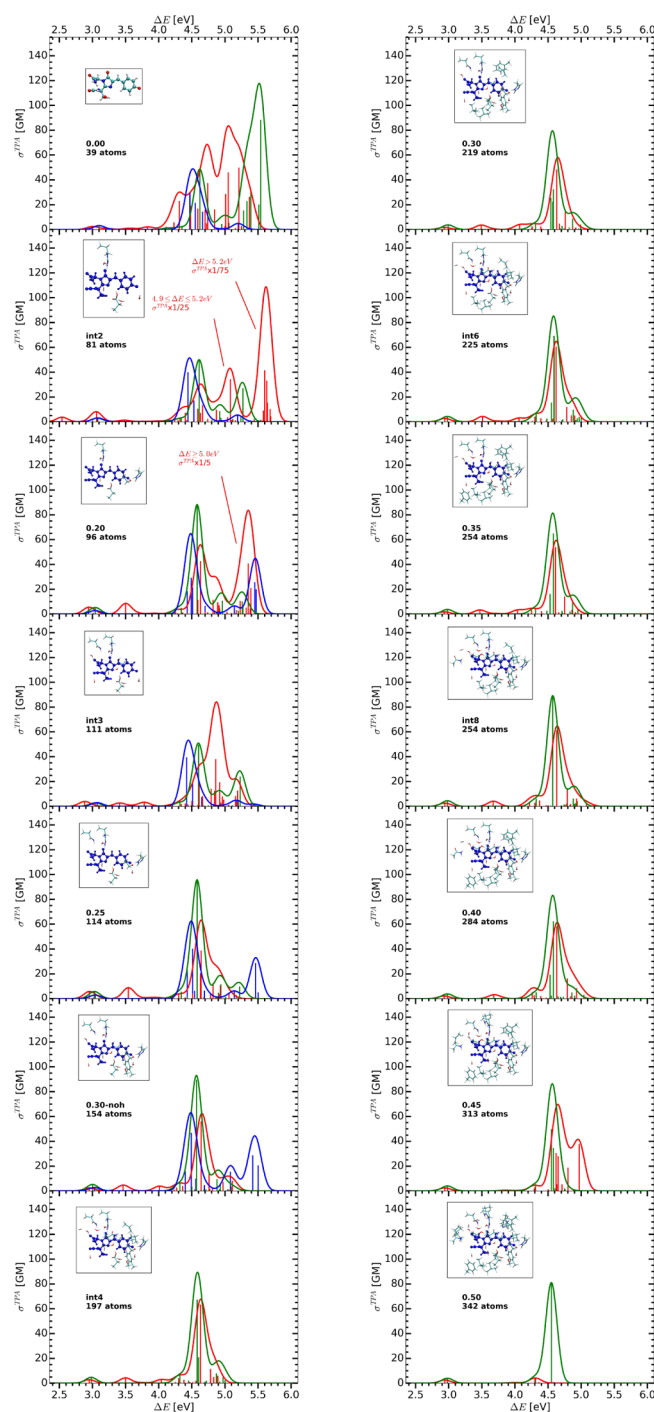

Figure 7. TPA spectra for chosen QM subsystems of avGFP-a models. The NE, EE, and PE results are given in red, green, and blue, respectively. For int2/NE and 0.20/NE models, the σ^{TPA} was rescaled in the chosen ΔE range as shown on plot to fit it.

OPA intensity. If Y203 is not part of the QM subsystem (0.00, int2, int3, 0.20, 0.25, and 0.30; Figure S9 in the Supporting Information), we observe much stronger dependence of σ^{TPA} intensity on the QM region size than in avGFP-a and avGFP-n. This clearly suggests that the chromophore—Y203 interaction is not well described by electrostatic monopoles which is consistent with a relatively small contribution of electrostatic interactions in benzene dimers.[93] In the case of the QM clusters including Y203, we observe that σ^{TPA} is usually in the range of 12−14 GM (NE) and 17−20 GM (EE; Tables S16−

S17 in the Supporting Information). As we found for OPA intensity, including more hydrophobic residues may damp TPA intensity (Figure 8). For instance, 0.35-noh has larger
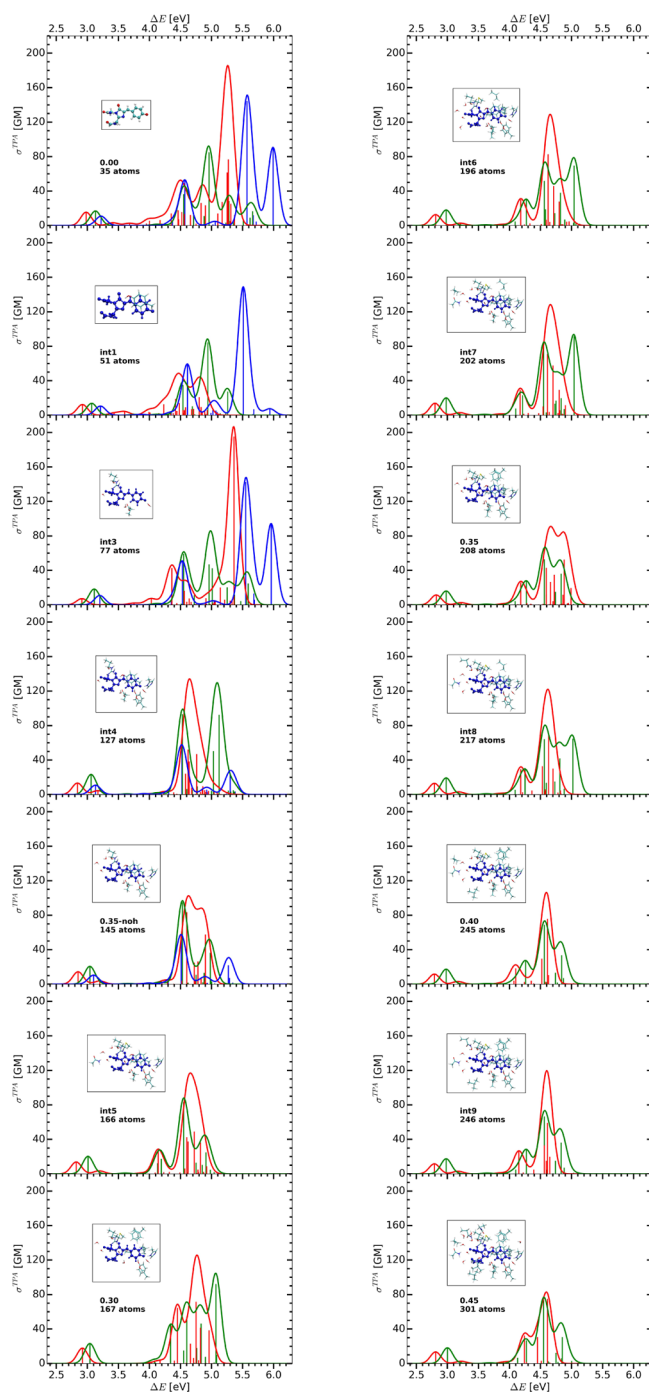

Figure 8. TPA spectra for chosen QM subsystems of Citrine models. The NE, EE, and PE results are given in red, green and blue, respectively.

σ^{TPA} than 0.35 by a few GM, but int5 and int7 have virtually the same σ^{TPA} although the latter is composed of three more hydrophobic a.a. residues.

It is quite astonishing that at the TDDFT/PE level of theory the TPA intensity for the $S_0 \rightarrow S_1$ transition virtually does not change with a QM cluster composition for any of investigated FPs (Figures 6−8). In the case of avGFP-n, the σ^{TPA} increases

from 0.7 to 1.1 GM for 0.00/PE and 0.30-noh/PE models, respectively, but reverts back to 0.9 GM for int4/PE. This picture is quite similar for the models of avGFP-a and Citrine, where the difference between the largest and smallest σ^{TPA} is 0.4 and 0.7 GM, respectively. When neutral/anionic GFP chromophore is extended by F64 and V68 a.a. residues, the σ^{TPA} decreases/increases by 0.6−1.4 GM/0.6−0.9 GM (depending on the protein conformation).[92] Thus, somehow a greater impact of the QM region size on the σ^{TPA} value was observed than using our protocol. Nevertheless, considering the absolute values of TPA cross section of 1.1−4.9 GM for the "small chromophore" therein (our 0.00/PE model), our observation of the minor QM region size impact on σ^{TPA} is in agreement with previous work.[92]

3.2.2. $S_0 \rightarrow S_n$ Transitions. 3.2.2.1. avGFP-n. According to Figure 6, there are distinctively bright TPA bands at *ca.* 4.2−4.7 eV for avGFP-n. This is also the case for excitations above 5.0 eV but for many QM clusters, we did not reach excited states responsible for creating this band. Hence, we focus on the lower-lying band unless stated otherwise. The TPA band is created by two virtually overlapping excited states S_3 and S_4 in the 0.00/NE model (Figure 6 and Table S10 in the Supporting Information). There is also a weak band at *ca.* 4.15−4.20 eV with σ^{TPA} equal 4.1 GM. For larger QM subsystems and NE, we observe two distinct TPA bands at 4.3 and 4.6 eV. As the QM subsystem size increases, their relative intensity changes rather nonlinearly (Figures 6 and S3 in the Supporting Information). For int8/NE−0.50/NE models consiting of at least 242 atoms, the lower-lying TPA band is always somehow weaker. These are very large QM subsystems for TPA spectrum calculations with $S_0 \rightarrow S_n$ transitions included. We find a similar feature for a smaller 0.35-noh/NE model (Figure S3 in the Supporting Information). Furthermore, one may observe a much weaker but distinct TPA band at *ca.* 3.8 eV in some NE models. It is dominated by $\pi_{H148} \rightarrow \pi^*$ transition and present in all models containing the H148 residue in the QM subsystem. It is not seen as a separate band in int3/NE and 0.25/NE systems because it is blue-shifted to 4.2 eV and gains σ^{TPA} of *ca.* 11 GM. As the QM subsystem size increases to 0.30-noh/NE and 0.35-noh/NE, the $\pi_{H148} \rightarrow \pi^*$ excited state is red-shifted to 3.8 eV and σ^{TPA} is *ca.* 6 GM (Table S10 in the Supporting Information). The main difference between int3 or 0.25 and 0.30-noh QM regions is the presence of S205 and more water molecules in the latter (Figure 2a and Tables S1− S3 in the Supporting Information). This leads to a complete h-bonds network between the chromophore and E222. Starting from int5/NE, which contains multiple hydrophobic residues compared to the above-mentioned QM subsystems, the discussed excited state becomes much dimmer for a TPA process with σ^{TPA} not exceeding 1.8 GM.

On the other hand, the qualitative features of the converged TPA spectrum in the discussed excitation energies range are already found for 0.00/EE (Figure 6). Even more, with the PE applied, there are small quantitative changes in the TPA spectrum features for two bands near 4.25−4.55 eV as the QM region is extended from 0.00 to int4 (Figures 6 and S5 in the Supporting Information). They are created by two excited states (S_2 and S_3). For the former, the ΔE, σ^{TPA}, and f values change in the range 4.23−4.26 eV, 6.9−9.4 GM, and 0.019− 0.024, respectively. For the latter, the range is 4.54−4.60 eV, 52.1−55.6 GM, and 0.033−0.038, respectively (Table S12 in the Supporting Information). The impact of the QM subsystem size in avGFP-n models is thus much smaller

within PE approximation as compared to NE and EE (see further text) cases. This clearly shows that neglecting interactions with the MM subsystem requires larger QM subsystems as also pointed out by others for the OPA spectrum of EGFP with an anionic chromophore.[35]

Nevertheless, within the EE model, the TPA spectrum is still affected by the QM subsystem size. As can be seen in Figure 6, the bright TPA band in the 4.2−4.7 eV range of excitation energies is usually created by three excited states (see also Table S11 in the Supporting Information). In int3/EE, 0.30-noh/EE, int4/EE, or 0.35-noh/EE models, we observe a clear split into four excited states. This can be attributed to a lack of or too small number of hydrophobic a.a. residues in these QM clusters. Besides, in the case of the int3/EE system which is the smallest one containing H148 residue, we observe a TPA bright (σ^{TPA} equal 34.1 GM) S_2 excited state (Table S11 in the Supporting Information) which is dominated by the $\pi_{H148} \rightarrow \pi^*$ transition. It is the main contributor to the TPA band at 4.3 eV (Figure 6). However, it is sufficient to include T62 and L220 (Figure 2a) in the 0.25 QM subsystem to make this excited state (S_4 in 0.25/EE model) much dimmer (σ^{TPA} equal 5.9 GM). It somehow regains a TPA intensity of 10.8−12.9 GM in the case of 0.30-noh/EE, int4/EE, and 0.35-noh/EE models but again becomes dimmer (2.9 GM) when the int5/EE model is used (Figure 6). It systematically loses TPA intensity as the QM subsystem size increases further on. Because 0.30-noh/EE, int4/EE, and 0.35-noh/EE systems do not contain hydrophobic a.a. residues, it is clear that they are required to suppress some two-photon excitations. More precisely, the L220 residue (see Tables S1−S3 in the Supporitng Information for QM clusters composition) seems to be mainly responsible for making the $\pi_{H148} \rightarrow \pi^*$ transition dimmer in the TPA process. We believe that one can assume the TPA band to be converged starting from int5/EE (Figure 6). ΔE does not change by more than 0.03 eV, and usually by 0.00−0.01 eV for the following QM subsystems (Table S11 in the Supporting Information). The σ^{TPA} value is 32−41 and 77−88 GM for the lowest and the highest excited states creating this TPA band, respectively. In the case of the middle excited state, we find its σ^{TPA} value to be ca. 20 GM if F165 belongs to the QM subsystem and ca. 30 GM otherwise. Once again, we observe importance of this a.a. residue in the shaping TPA spectrum of avGFP-n. As a final remark, the TPA band above 5.0 eV is formed by excitations characterized by transitions into undefined or π_{cap} MOs. Thus, our QM subsystems are not large enough to reliably describe this TPA band. This is somehow improved when the PE is applied. The TPA band at $\Delta E > 5.0$ eV, as shown in Figure S9 in the Supporting Information, is created by a few excited states. Some of them are characterized by a charge transfer from the E222 a.a. residue to the chromophore while others display $\pi \rightarrow \pi^*$ character with MOs localized on the chromophore. It is interesting to note that for the QM subsystems containing H148 a.a. residue (int3, 0.25, 0.30-noh and int4 combined with PE), there is a distinctive band at 5.0 eV (it is blue-shifted to ca 5.2 eV for int4/PE model). It is created by excited states with a significant contribution from the $\pi \rightarrow \pi^*$ transition located on the H148 a.a. residue. Excited states characterized by a similar transition were also found with EE but they were significantly redshifted and became dimmer and dimmer for the TPA process as the QM subsystem size increased. In an upcoming paper, we argue that the TPA bright excited state involving

H148 can be an artifact in PE calculations because of a lack of Pauli repulsion.

3.2.2.2. avGFP-a. In the case of 0.00/NE model of avGFP-a, we observe an extensive TPA band ranging from 4.0 to 5.5 eV (Figure 7). It is created by ca. dozen of excited states (Table S13 in the Supporting Information). The QM models extended by addition of R96, E222 (int1/NE), and also four water molecules (int2/NE), reveal extremely large σ^{TPA} values for transitions into excited states within 5.0−5.5 eV. This may be related to a resonant enhancement[9] because of the presence of artificious excited states lying below the brightest excited state for the OPA process (usually S_1). Their excitation energies are 2.30−2.93 eV (int1/NE) and 2.54 eV (int2/NE), and excited states displaying enormous σ^{TPA} value have often ca. two times larger ΔE in the range of 5.42−5.60 and 5.08 or 5.44−5.69 eV, respectively. These data clearly show shortcomings of too small QM subsystems. This picture is somehow improved in 0.20/NE (int1/NE + H148, T203 and two water molecules h-bonded to the chromophore—Figure 2b) model: we do not observe excited states with artificially low excitation energy and the brightest TPA transitions are more localized on the chromophore and its surroundings (in particular R96, H148 and T203). Also including EE in our model, leads to a well-defined TPA band near 4.6 eV created by a single $\pi \rightarrow \pi^*$ transition for all models discussed so far. In the case of the 0.20/EE model and models with larger QM regions (except int3), we observe an almost two-fold increase in the σ^{TPA} value compared to smaller QM subsystems and int3/EE (Figure 7). Apparently, this is because of the presence of H148 and T203 in the 0.20 QM region (Tables S4−S6 in the Supporting Information). Taking into consideration a more pronounced TPA process in the EE case compared to NE near 4.5 eV, we conclude that a.a. residues h-bonded to the chromophore work together with electrostatic interactions to enhance TPA process in avGFP-a. In the EE case, the discussed transition is split into two excited states for most of the QM subsystems starting from int4. They exhibit σ^{TPA} in range of 8.3−20.8 and 62.2−73.5 GM and are separated by 0.02−0.07 eV (Table S14 in the Supporting Information). In the case of the 0.45/EE model, the excitation intensity transfer leads to more equal σ^{TPA} values while in int8/EE and 0.50/EE, we do not observe split excited states for this band (Figure 7). We also observe an excited state of $\pi \rightarrow \pi^*$ character at ca. 4.3 eV. It contributes to the TPA band only slightly as its σ^{TPA} value does not exceed 6.0 GM and is almost immune to the QM subsystem size (Table S14 in the Supporting Information). We conclude that the QM subsystem composition must be carefully chosen to obtain a qualitatively and quantitatively correct TPA band. We cannot strictly say that we reached quantitatively converged TPA spectrum in discussed range of excitation energies even with an extensive 0.50 QM subsystem consisting of 342 atoms. Whether the discussed TPA band consists of one or two excited states is dictated by a complex game of interactions within the QM subsystem and between QM and MM subsystems. This is supported by the fact that for QM regions significantly differing in size and composition (Tables S4−S6 in the Supporting Information): 0.25, 0.30-noh, int8, and 0.50, we observe one excited state being responsible for discussed TPA band (Figure 7).

For avGFP-a models with the PE included, the part of the TPA spectrum at ca. 4.5 eV (Figure 7) is created by three excited states if the H148 a.a. residue is not part of the QM subsystem. As the QM region is extended starting from 0.00

through int1, int2 to int3 (all QM regions without H148 and combined with PE), we observe that the S_2 excited state becomes slightly red-shifted and gains larger σ^{TPA} at the cost of the S_4 state (Figure 7 and Table S15 in the Supporting Information). Overall, as the a.a. residues h-bonded to the chromophore are included in the QM region, the TPA band at *ca.* 4.5 eV is only slightly enhanced. This picture changes somehow for the 0.20/PE, 0.25/PE, and 0.30-noh/PE models with the H148 included in the QM region (Figure 7). Then, we observe four excited states forming the TPA band at 4.5 eV. This is due to an extra excited state (S_3 in 0.20/PE, S_4 in 0.25/PE, and 0.30-noh/PE—see Table S15 in the Supporting Information) characterized by a transition from π MO localized on the chromophore into an undefined MO on H148. It is quite bright in the 0.20/PE model (29.2 GM) but becomes dimmer (6.1−9.8 GM) for larger QM subsystems. The presence of H148 in the QM region leads also to a bright TPA band peaking at *ca.* 5.5 eV and is in fact created by one excited state of charge transfer character from the chromophore to H148 and one excitation localized on the latter a.a. residue.

3.2.2.3. Citrine. In the case of small QM subsystems of Citrine FP not containing Y203 (0.00, int2, int3, 0.20, 0.25), we have similar conclusions as for avGFP-a FP models of similar size regarding σ^{TPA} value—QM subsystem size relationship. We thus focus on TPA spectra for QM clusters containing Y203 (Tables S7−S9 in the Supporting Information). As it was the case for the OPA spectrum, including Y203 in the QM region (Figure 2c) visibly changes the TPA spectrum unless PE is applied (Figure 8). The brightest TPA peaks for 0.00/PE and int1/PE models are hardly distinguishable.

In the NE case, we find pretty weak TPA band (σ^{TPA} of 2.5−4.1 GM) created by two excited states at 3.43 and 3.60 eV (Figure 8 and Table S16 in the Supporting Information). The former is characterized by a $\pi \rightarrow \pi^*$ transition while in the latter, we observe a transition into undefined MO. The latter is not observed while the former is red-shifted to *ca.* 3.2 eV if larger QM subsystems are used. By applying EE, this TPA band is hardly detected, as shown in Figure 8, at *ca.* 3.6−3.7 eV because of the σ^{TPA} value not exceeding 1.0 GM regardless of the QM subsystem size (Table S17 in the Supporting Information). Also within PE approximation, we find a TPA dim band at 3.9−4.0 eV (Table S18 and Figure S11 in the Supporting Information). It is created by a S_2 excited state of charge-transfer character from Y203 to the chromophore. The extensive TPA band at 4.0−5.0 eV of the int1/NE model (chromophore + Y203) is created by multiple excited states often of charge-transfer character between chromophore and Y203 in either way as found previously by Beerepoot et al.[44] The part of the Citrine TPA spectrum above 4.0 eV significantly when the QM subsystem is extended (Figures 8 and S7 in the Supporting Information). In the range of 4.1−4.3 eV, we observe an excitation peak in both NE and EE cases which is usually due to excitation of the free electron pair of M69 (Figure 2c) sulfur atom to π MO localized on the chromophore. As the QM subsystem size increases, or PE is applied, it does not disappear thus we think this is not an artifact. However, we observe a strong dependence of the σ^{TPA} value on the QM cluster size for 0.40/NE, int9/NE, and 0.45/NE models but not when we use EE (Figure 8). The nature of the TPA peak at 4.5−4.6 eV (NE) depends more strongly on the QM cluster size. It is created by one or two excited states

dominated by $\pi \rightarrow \pi^*$ transition or an excitation from the $\pi + \pi_{Y203} \rightarrow \pi^*$ MO into π^*_{cap} or an undefined MO. Quite surprisingly,[44] the excited states displaying charge-transfer between Y203 and the chromophore do not significantly contribute to strong TPA in the QM subsystems larger than int1/NE. In most cases, we observe one TPA band with σ^{TPA} in the range of 90−130 GM (Figures 8 and S7 in the Supporting Information). This picture is somehow different for the 0.35/NE system presumably because of the stronger TPA process for excited states "on the edge" (S_{16} and S_{26}—Table S16 in the Supporting Information) of the band than in other models. As a consequence, we may distinguish two slightly separated bands (Figure 8). However, if we use a more complete model of FP by applying EE, we in fact observe an even larger separation of two distinct TPA bands (Figures 8 and S7 in the Supporting Information). The TPA peak at *ca.* 4.5 eV has the intensity over 90 GM for int4/EE and 0.35-noh/EE models which do not contain hydrophobic a.a. residues in the QM subsystems. This peak is created by the one excited state (Table S17 in the Supporting Information). For larger QM regions, except int5, this peak becomes somehow dimmer (70−80 GM). This is presumably due to the M69 presence which "steals" the TPA process intensity as it is involved in lower-lying transition as we have already discussed. Furthermore, in the case of 0.30/EE (Y203 is not part of the QM region), int6/EE, 0.35/EE, and int8/EE models (Figure 8), we observe two slightly separated excited states with ΔE in the range of 4.55−4.60 eV and σ^{TPA} in the range of 11.7−63.9 GM creating the discussed TPA band. When other QM subsystems (int5, int7, int9, 0.40−0.50) are combined with EE, we again find one distinctively TPA bright state (*ca.* 70−80 GM) while the other has a σ^{TPA} smaller than 5 GM (Table S17 in the Supporting Information). It is hard to point to a single QM cluster composition (Tables S7−S9 in the Supporting Information) descriptor responsible for this qualitative feature of TPA band. However, it seems that the QM subsystems for which single excited state is detected are rich in hydrophobic a.a. residues, except for int5 and int7 clusters. The higher-lying TPA band ($\Delta E > 4.5$ eV) consists of excited states having various character. These are often transitions from π-type MO delocalized over the chromophore and Y203 into an undefined orbital, as well as $n_{ph} \rightarrow \pi^*$. n_{ph} denotes MO describing a free electron pair of the chromophore's phenyl oxygen which is actually spread over H148 and Y145 residues showing their involvement in shaping the TPA spectrum. Starting from the 0.40/EE model (245 atoms), we can assume that the discussed part of the TPA spectrum is well converged with respect to the QM subsystem size (Figure S7 in the Supporting Information). In the case of 0.50/EE, we do not find the highest energy peak peak because we were not able to calculate enough excited states. We find a qualitatively similar TPA spectrum for a much smaller 0.35/EE model (208 atoms).

The TPA peaks at 4.6 and 5.0 eV as predicted by the int1/PE model become red-shifted by about 0.1 eV when int4/PE and 0.35-noh/PE models are used (Figure 8), which are about 2.5−3 times larger in terms of the QM region size. Simultaneously, the TPA intensity changes negligibly—by less than 2 GM for the lower-lying peak (Table S18 in the Supporting Information). The peak at *ca.* 5.3 eV in the int4/PE and 0.35-noh/PE models is again created by a transition localized on the H148 a.a. residue. On the other hand, the TPA band at 5.5 eV, predicted by int1/PE model, is not seen in the

spectra of the remaining models, as we did not reach the required excitation energies region in the latters.

4. CONCLUSIONS

We systematically investigated impact of QM subsystem size and composition on calculated OPA and TPA spectra of three FPs. The brightest OPA band is qualitatively well described by QM region consisting of chromophore only (0.00), but it becomes red-shifted by up to 0.15 eV until ΔE does not change with increasing QM subsystem size further. This is the case for EE, while without any QM and MM subsystems interactions, the redshift is even larger. The OPA intensity as measured by oscillator strength converges somehow slower than ΔE, as even for QM subsystems consisting of more than ca. 250 atoms, f value slowly decreases with the QM subsystem size. It was found that residues h-bonded to the chromophore enhance OPA process. On the other hand, hydrophobic a.a. residues cannot be neglected when choosing QM subsystem because they account for a few hundredths electronvolts of a redshift as well as they damp OPA intensity. Moreover, the hydrophobic residues are important for the suppressing TPA process to spurious excited states involving charge-transfer between the chromophore and a.a. residues, for example, H148.

The TPA spectrum is more challenging to simulate than OPA spectrum in cases where NE or EE are in action. According to our results, it is evident that the converged TPA spectrum requires more extensive QM cluster than the OPA spectrum. The minimal QM subsystem consisting of the chromophore only, leads to neither qualitatively nor quantitatively correct TPA spectrum in the >4.0 eV region of excitation energies. Usually, we find QM subsystems consisting of ca. 220 atoms for which OPA and TPA spectra contain qualitative features of those obtained with much larger QM subsystems (>300 atoms). This is especially the case when EE is applied, which allows to use smaller QM subsystem as also suggested by others for OPA.[35]

When the MM subsystem is approximated by the PE, we observe a rather small impact of the QM subsystem size on both OPA and TPA spectra. This is true for the $S_0 \rightarrow S_1$ and higher transitions. As a result significantly smaller QM regions, consisting of ca. 80–100 atoms, than in the EE scheme, can be utilized to obtain converged OPA and TPA properties.

Based on our results, we propose a following algorithm to decide about the QM subsystem size and composition for OPA and TPA spectra calculations in FPs. First, all a.a. residues and water molecules within 0.30 nm (chromophore made of SYG triplet as in avGFP-n and avGFP-a) or 0.35 nm (chromophore made of GYG triplet as in Citrine) from the chromophore should be part of the QM subsystem if the EE is applied. MD simulation run snapshots will provide a structural content of the protein appearing most frequently in a given radius. A visual inspection should be made to find a.a. residues and water molecules that may happen to be further from the chromophore but are involved in an extensive h-bond network. When the QM subsystem size precludes TPA calculations, one should consider removing some hydrophobic a.a. residues. We warn, however, that the effect of changing the QM subsystem composition on absorption spectra must be always carefully checked. In the case of avGFP-n, but not avGFP-a and Citrine, we find that F165 may be important for the TPA spectrum fine-tuning. In the case of Citrine, Y203 must be obviously included in the QM subsystem together with M69.

Presumably, all methionine and cysteine residues in 0.30–0.35 nm radius from the chromophore should be described with QM methodology. On the other hand, if PE is applied the chromophore alone in the QM region is already enough to obtain qualitatively converged OPA and TPA spectra although we advice to include R96, E222, and water molecules h-bonded to the chromophore, which are poorly described by PE localized multipoles and polarizabilities,[22] as well as Y203 for the YFP models. The computational cost is still affordable and the results are more reliable especially in the case of the avGFP-a model. One must be aware though that to draw quantitative conclusions about the role of a specific a.a. in shaping OPA and TPA spectra a more reliable phase-space sampling may be required.

We believe that our work is a firm source of data for rational choice of QM region composition in OPA and TPA spectra calculations in FPs. Our guidance can be also used to build models of other photoactive proteins for absorption spectra calculations. It should be useful when attempting simulation of absorption spectra using multiple snapshots from MD run since this requires reliable results at the lowest cost possible.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jctc.0c00602.

> Technical details of classical and hybrid MD simulations protocols, composition of QM subsystems, results of OPA and TPA simulations, simulated OPA and TPA spectra, comparison of OPA and TPA results obtained with BHandHLYP and CAM-B3LYP, and impact of number of water molecules included in PE model on OPA and TPA results (PDF)

> Cartesian coordinates (in Gromacs format) of avGFP-a, avGFP-n and Citrine fluorescent protein structures utilized for classical molecular dynamics simulations (*-crystall.gro) as well as for absorption spectra calculations (*-md-qmmm.gro) (ZIP)

■ AUTHOR INFORMATION

Corresponding Author

Tadeusz Andruniów − *Advanced Materials Engineering and Modelling Group, Wroclaw University of Science and Technology, 50-370 Wroclaw, Poland;* ⓞ orcid.org/0000-0002-2215-5287; Email: tadeusz.andruniow@pwr.edu.pl

Author

Dawid Grabarek − *Advanced Materials Engineering and Modelling Group, Wroclaw University of Science and Technology, 50-370 Wroclaw, Poland;* ⓞ orcid.org/0000-0001-5815-9987

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.jctc.0c00602

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

chromophores. This work was supported by National Science Centre, Poland (grant no. 2016/21/N/ST4/00004). Calculations were performed at the Wroclaw Centre for Networking and Supercomputing (WCSS).

■ REFERENCES

(1) Shimomura, O.; Johnson, F. H.; Saiga, Y. Extraction, Purification and Properties of Aequorin, a Bioluminescent Protein from the Luminous Hydromedusan, Aequorea. *J. Cell. Comp. Physiol.* **1962**, *59*, 223−239.

(2) Morise, H.; Shimomura, O.; Johnson, F. H.; Winant, J. Intermolecular energy transfer in the bioluminescent system of Aequorea. *Biochemistry* **1974**, *13*, 2656−2662.

(3) Stepanenko, O.; Verkhusha, V.; Kuznetsova, I. M.; Uversky, V.; Turoverov, K. Fluorescent Proteins as Biomarkers and Biosensors: Throwing Color Lights on Molecular and Cellular Processes. *Curr. Protein Pept. Sci.* **2008**, *9*, 338−369.

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

(5) Chudakov, D. M.; Matz, M. V.; Lukyanov, S.; Lukyanov, K. A. Fluorescent Proteins and Their Applications in Imaging Living Cells and Tissues. *Physiol. Rev.* **2010**, *90*, 1103−1163.

(6) Zipfel, W. R.; Williams, R. M.; Webb, W. W. Nonlinear magic: multiphoton microscopy in the biosciences. *Nat. Biotechnol.* **2003**, *21*, 1369−1377.

(7) Benninger, R. K. P.; Piston, D. W. Two-photon excitation microscopy for the study of living cells and tissues. *Curr. Protoc. Cell Biol.* **2013**, *59*, 4.11.1−4.11.24.

(8) Adam, V.; Mizuno, H.; Grichine, A.; Hotta, J.-i.; Yamagata, Y.; Moeyaert, B.; Nienhaus, G. U.; Miyawaki, A.; Bourgeois, D.; Hofkens, J. Data storage based on photochromic and photoconvertible fluorescent proteins. *J. Biotechnol.* **2010**, *149*, 289−298.

(9) Drobizhev, M.; Makarov, N. S.; Tillo, S. E.; Hughes, T. E.; Rebane, A. Two-photon absorption properties of fluorescent proteins. *Nat. Methods* **2011**, *8*, 393−399.

(10) Molina, R. S.; Tran, T. M.; Campbell, R. E.; Lambert, G. G.; Salih, A.; Shaner, N. C.; Hughes, T. E.; Drobizhev, M. Blue-Shifted Green Fluorescent Protein Homologues Are Brighter than Enhanced Green Fluorescent Protein under Two-Photon Excitation. *J. Phys. Chem. Lett.* **2017**, *8*, 2548−2554.

(11) Nifosì, R.; Luo, Y. Predictions of Novel Two-Photon Absorption Bands in Fluorescent Proteins. *J. Phys. Chem. B* **2007**, *111*, 14043−14050.

(12) Drobizhev, M.; Tillo, S.; Makarov, N. S.; Hughes, T. E.; Rebane, A. Absolute Two-Photon Absorption Spectra and Two-Photon Brightness of Orange and Red Fluorescent Proteins. *J. Phys. Chem. B* **2009**, *113*, 855−859.

(13) Oulianov, D. A.; Tomov, I. V.; Dvornikov, A. S.; Rentzepis, P. M. Observations on the measurements of two-photon absorption cross-section. *Opt. Commun.* **2001**, *191*, 235−243.

(14) Cranfill, P. J.; Sell, B. R.; Baird, M. A.; Allen, J. R.; Lavagnino, Z.; de Gruiter, H. M.; Davidson, M. W.; Piston, D. W. Quantitative assessment of fluorescent proteins. *Nat. Methods* **2016**, *13*, 557−562.

(15) Hosoi, H.; Yamaguchi, S.; Mizuno, H.; Miyawaki, A.; Tahara, T. Hidden electronic excited state of enhanced green fluorescent protein. *J. Phys. Chem. B* **2008**, *112*, 2761−2763.

(16) Hashimoto, H.; Isobe, K.; Suda, A.; Kannari, F.; Kawano, H.; Mizuno, H.; Miyawaki, A.; Midorikawa, K. Measurement of two-photon excitation spectra of fluorescent proteins with nonlinear Fourier-transform spectroscopy. *Appl. Opt.* **2010**, *49*, 3323−3329.

(17) Blab, G. A.; Lommerse, P. H. M.; Cognet, L.; Harms, G. S.; Schmidt, T. Two-photon excitation action cross-sections of the autofluorescent proteins. *Chem. Phys. Lett.* **2001**, *350*, 71−77.

(18) Heikal, A. A.; Hess, S. T.; Webb, W. W.; Tsien, R. Y.; Webb, W. W. Multiphoton molecular spectroscopy and excited-state dynamics of enhanced green fluorescnt protein (EGFP): acid−base specificity. *Chem. Phys.* **2001**, *274*, 37−45.

(19) Kawano, H.; Kogure, T.; Abe, Y.; Mizuno, H.; Miyawaki, A. Two-photon dual-color imaging using fluorescent proteins. *Nat. Methods* **2008**, *5*, 373−374.

(20) Rizzo, M. A.; Springer, G.; Segawa, K.; Zipfel, W. R.; Piston, D. W. Optimization of pairings and detection conditions for measurements of FRET between cyan and yellow fluorescent proteins. *Microsc. Microanal.* **2006**, *12*, 238−254.

(21) List, N. H.; Olsen, J. M. H.; Jensen, H. J. A.; Steindal, A. H.; Kongsted, J. Molecular-Level Insight into the Spectral Tuning Mechanism of the DsRed Chromophore. *J. Phys. Chem. Lett.* **2012**, *3*, 3513−3521.

(22) Steindal, A. H.; Olsen, J. M. H.; Ruud, K.; Frediani, L.; Kongsted, J. A combined quantum mechanics/molecular mechanics study of the one- and two-photon absorption in the green fluorescent protein. *Phys. Chem. Chem. Phys.* **2012**, *14*, 5440−5451.

(23) Drobizhev, M.; Tillo, S.; Makarov, N. S.; Hughes, T. E.; Rebane, A. Color Hues in Red Fluorescent Proteins Are Due to Internal Quadratic Stark Effect. *J. Phys. Chem. B* **2009**, *113*, 12860−12864.

(24) Grabarek, D.; Andruniów, T. Illuminating the origins of two-photon absorption properties in fluorescent protein chromophores. *Int. J. Quantum Chem.* **2020**, *120*, No. e26086.

(25) Salem, M. A.; Twelves, I.; Brown, A. Prediction of two-photon absorption enhancement in red fluorescent protein chromophores made from non-canonical amino acids. *Phys. Chem. Chem. Phys.* **2016**, *18*, 24408−24416.

(26) Alam, M. M.; Chattopadhyaya, M.; Chakrabarti, S.; Ruud, K. Chemical Control of Channel Interference in Two-Photon Absorption Processes. *Acc. Chem. Res.* **2014**, *47*, 1604−1612.

(27) Cronstrand, P.; Luo, Y.; Ågren, H. Generalized few-state models for two-photon absorption of conjugated molecules. *Chem. Phys. Lett.* **2002**, *352*, 262−269.

(28) Cronstrand, P.; Luo, Y.; Ågren, H. Effects of dipole alignment and channel interference on two-photon absorption cross sections of two-dimensional charge-transfer systems. *J. Chem. Phys.* **2002**, *117*, 11102−11106.

(29) Warshel, A.; Levitt, M. Theoretical studies of enzymatic reactions: dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J. Mol. Biol.* **1976**, *103*, 227−249.

(30) Nifosì, R.; Mennucci, B.; Filippi, C. The key to the yellow-to-cyan tuning in the green fluorescent protein family is polarisation. *Phys. Chem. Chem. Phys.* **2019**, *21*, 18988−18998.

(31) Daday, C.; Curutchet, C.; Sinicropi, A.; Mennucci, B.; Filippi, C. Chromophore-Protein Coupling beyond Nonpolarizable Models: Understanding Absorption in Green Fluorescent Protein. *J. Chem. Theory Comput.* **2015**, *11*, 4825−4839.

(32) Beerepoot, M. T. P.; Steindal, A. H.; Kongsted, J.; Brandsdal, B. O.; Frediani, L.; Ruud, K.; Olsen, J. M. H. A Polarizable Embedding DFT Study of One-Photon Absorption in Fluorescent Proteins. *Phys. Chem. Chem. Phys.* **2013**, *15*, 4735−4743.

(33) Schwabe, T.; Beerepoot, M. T. P.; Olsen, J. M. H.; Kongsted, J. Analysis of computational models for an accurate study of electronic excitations in GFP. *Phys. Chem. Chem. Phys.* **2015**, *17*, 2582−2588.

(34) Beerepoot, M. T. P.; Steindal, A. H.; Ruud, K.; Olsen, J. M. H.; Kongsted, J. Convergence of environment polarization effects in multiscale modeling of excitation energies. *Comput. Theor. Chem.* **2014**, *1040−1041*, 304−311.

(35) Nåbo, L. J.; Olsen, J. M. H.; Martínez, T. J.; Kongsted, J. The Quality of the Embedding Potential is Decisive for Minimal Quantum Region Size in Embedding Calculations: The Case of the Green Fluorescent Protein. *J. Chem. Theory Comput.* **2017**, *13*, 6230−6236.

(36) Griesbeck, O.; Baird, G. S.; Campbell, R. E.; Zacharias, D. A.; Tsien, R. Y. Reducing the environmental sensitivity of yellow fluorescent protein. Mechanism and applications. *J. Biol. Chem.* **2001**, *276*, 29188−29194.

(37) Bublitz, G.; King, B. A.; Boxer, S. G. Electronic Structure of the Chromophore in Green Fluorescent Protein (GFP). *J. Am. Chem. Soc.* **1998**, *120*, 9370−9371.

(38) Chattoraj, M.; King, B. A.; Bublitz, G. U.; Boxer, S. G. Ultra-fast excited state dynamics in green fluorescent protein: multiple states and proton transfer. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 8362−8367.

(39) Brejc, K.; Sixma, T. K.; Kitts, P. A.; Kain, S. R.; Tsien, R. Y.; Ormö, M.; Remington, S. J. Structural basis for dual excitation and photoisomerization of the Aequorea victoria green fluorescent protein. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 2306−2311.

(40) Palm, G. J.; Zdanov, A.; Gaitanaris, G. A.; Stauber, R.; Pavlakis, G. N.; Wlodawer, A. The structural basis for spectral variations in green fluorescent protein. *Nat. Struct. Biol.* **1997**, *4*, 361−365.

(41) van Thor, J. J.; Pierik, A. J.; Nugteren-Roodzant, I.; Xie, A.; Hellingwerf, K. J. Characterization of the Photoconversion of Green Fluorescent Protein with FTIR Spectroscopy. *Biochemistry* **1998**, *37*, 16915−16921.

(42) Ormö, M.; Cubitt, A. B.; Kallio, K.; Gross, L. A.; Tsien, R. Y.; Remington, S. J. Crystal structure of the Aequorea victoria green fluorescent protein. *Science* **1996**, *273*, 1392−1395.

(43) Wachter, R. M.; Elsliger, M.-A.; Kallio, K.; Hanson, G. T.; Remington, S. J. Structural basis of spectral shifts in the yellow-emission variants of green fluorescent protein. *Structure* **1998**, *6*, 1267−1277.

(44) Beerepoot, M. T. P.; Friese, D. H.; Ruud, K. Intermolecular charge transfer enhances two-photon absorption in yellow fluorescent protein. *Phys. Chem. Chem. Phys.* **2014**, *16*, 5958−5964.

(45) Becke, A. D. A new mixing of Hartree-Fock and local density-functional theories. *J. Chem. Phys.* **1993**, *98*, 1372−1377.

(46) Grabarek, D.; Andruniów, T. Assessment of Functionals for TDDFT Calculations of One- and Two-Photon Absorption Properties of Neutral and Anionic Fluorescent Proteins Chromophores. *J. Chem. Theory Comput.* **2019**, *15*, 490−508.

(47) Elsliger, M.-A.; Wachter, R. M.; Hanson, G. T.; Kallio, K.; Remington, S. J. Structural and spectral response of green fluorescent protein variants to changes in pH. *Biochemistry* **1999**, *38*, 5296−5301.

(48) Filippi, C.; Buda, F.; Guidoni, L.; Sinicropi, A. Bathochromic Shift in Green Fluorescent Protein: A Puzzle for QM/MM Approaches. *J. Chem. Theory Comput.* **2012**, *8*, 112−124.

(49) Dunbrack, R. L., Jr. Rotamer libraries in the 21st century. *Curr. Opin. Struct. Biol.* **2002**, *12*, 431−440.

(50) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. UCSF Chimera - a visualization system for exploratory research and analysis. *J. Comput. Chem.* **2004**, *25*, 1605−1612.

(51) 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 supercomputers. *SoftwareX* **2015**, *1−2*, 19−25.

(52) Takaba, K.; Tai, Y.; Eki, H.; Dao, H.-A.; Hanazono, Y.; Hasegawa, K.; Miki, K.; Takeda, K. Subatomic resolution X-ray structures of green fluorescent protein. *IUCrJ* **2019**, *6*, 387−400.

(53) Shibazaki, C.; Shimizu, R.; Kagotani, Y.; Ostermann, A.; Schrader, T. E.; Adachi, M. Direct Observation of the Protonation States in the Mutant green fluorescent Protein. *J. Phys. Chem. Lett.* **2020**, *11*, 492−496.

(54) Reuter, N.; Lin, H.; Thiel, W. Green Fluorescent Proteins: Empirical Force Field for the Neutral and Deprotonated Forms of the Chromophore. Molecular Dynamics Simulations of the Wild Type and S65T Mutant. *J. Phys. Chem. B* **2002**, *106*, 6310−6321.

(55) 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.

(56) Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins* **2010**, *78*, 1950−1958.

(57) Best, R. B.; Hummer, G. Optimized Molecular Dynamics Force Fields Applied to the Helix-Coil Transition of Polypeptides. *J. Phys. Chem. B* **2009**, *113*, 9004−9015.

(58) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of multiple AMBER force fields and development of improved protein backbone parameters. *Proteins* **2006**, *65*, 712−725.

(59) Storti, B.; Margheritis, E.; Abbandonato, G.; Domenichini, G.; Dreier, J.; Testa, I.; Garau, G.; Nifosì, R.; Bizzarri, R. Role of Gln222 in Photoswitching of Aequorea Fluorescent Proteins: A Twisting and H-Bonding Affair? *ACS Chem. Biol.* **2018**, *13*, 2082−2093.

(60) Hutter, J.; Iannuzzi, M.; Schiffmann, F.; VandeVondele, J. CP2K: atomistic simulations of condensed matter systems. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2014**, *4*, 15−25.

(61) Becke, A. D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A* **1988**, *38*, 3098−4000.

(62) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **1988**, *37*, 785−789.

(63) A development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007. *TURBOMOLE V7.3*; TURBOMOLE GmbH, since 2007; available from, 2018 http://www.turbomole.com.

(64) *Dalton, A Molecular Electronic Structure Program Release DALTON2018.0*, 2018, see http://daltonprogram.org.

(65) Aidas, K.; Angeli, C.; Bak, K. L.; Bakken, V.; Bast, R.; Boman, L.; Christiansen, O.; Cimiraglia, R.; Coriani, S.; Dahle, P.; Dalskov, E. K.; Ekström, U.; Enevoldsen, T.; Eriksen, J. J.; Ettenhuber, P.; Fernández, B.; Ferrighi, L.; Fliegl, H.; Frediani, L.; Hald, K.; Halkier, A.; Hättig, C.; Heiberg, H.; Helgaker, T.; Hennum, A. C.; Hettema, H.; Hjertenaes, E.; Høst, S.; Høyvik, I.-M.; Iozzi, M. F.; Jansík, B.; Jensen, H. J. A.; Jonsson, D.; Jørgensen, P.; Kauczor, J.; Kirpekar, S.; Kjaergaard, T.; Klopper, W.; Knecht, S.; Kobayashi, R.; Koch, H.; Kongsted, J.; Krapp, A.; Kristensen, K.; Ligabue, A.; Lutnaes, O. B.; Melo, J. I.; Mikkelsen, K. V.; Myhre, R. H.; Neiss, C.; Nielsen, C. B.; Norman, P.; Olsen, J.; Olsen, J. M. H.; Osted, A.; Packer, M. J.; Pawlowski, F.; Pedersen, T. B.; Provasi, P. F.; Reine, S.; Rinkevicius, Z.; Ruden, T. A.; Ruud, K.; Rybkin, V. V.; Sałek, P.; Samson, C. C. M.; de Merás, A. S.; Sauer, S. P. A.; Schimmelpfennig, B.; Sneskov, K.; Steindal, A. H.; Sylvester-Hvid, K. O.; Taylor, P. R.; Teale, A. M.; Tellgren, E. I.; Tew, D. P.; Thorvaldsen, A. J.; Thøgersen, L.; Vahtras, O.; Watson, M. A.; Wilson, D. J. D.; Ziolkowski, M.; Ågren, H.; Ågren, H. The Dalton quantum chemistry program system. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2014**, *4*, 269−284.

(66) Sałek, P.; Vahtras, O.; Helgaker, T.; Ågren, H. Density-functional theory of linear and nonlinear time-dependent molecular properties. *J. Chem. Phys.* **2002**, *117*, 9630−9645.

(67) Olsen, J.; Jo/rgensen, P. Linear and nonlinear response functions for an exact state and for an MCSCF state. *J. Chem. Phys.* **1985**, *82*, 3235−3264.

(68) Kendall, R. A.; Dunning, T. H., Jr.; Harrison, R. J. Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions. *J. Chem. Phys.* **1992**, *96*, 6796−6806.

(69) 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.

(70) A development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007. *TURBOMOLE V7.5*; TURBOMOLE GmbH, since 2007; available from, 2020 http://www.turbomole.com.

(71) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104.

(72) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* **2011**, *32*, 1456−1465.

(73) Goerigk, L.; Grimme, S. A thorough benchmark of density functional methods for general main group thermochemistry, kinetics, and noncovalent interactions. *Phys. Chem. Chem. Phys.* **2011**, *13*, 6670−6688.

(74) Eichkorn, K.; Treutler, O.; Öhm, H.; Häser, M.; Ahlrichs, R. Auxiliary basis sets to approximate Coulomb potentials. *Chem. Phys. Lett.* **1995**, *240*, 283−290.

(75) Eichkorn, K.; Treutler, O.; Öhm, H.; Häser, M.; Ahlrichs, R. Erratum: Auxiliary basis sets to approximate Coulomb potentials. *Chem. Phys. Lett.* **1995**, *242*, 652−660.

(76) Beerepoot, M. T. P.; Friese, D. H.; List, N. H.; Kongsted, J.; Ruud, K. Benchmarking two-photon absorption cross sections: performance of CC2 and CAM-B3LYP. *Phys. Chem. Chem. Phys.* **2015**, *17*, 19306−19314.

(77) Salem, M. A.; Brown, A. Two-Photon Absorption in Fluorescent Protein Chromophores: TDDFT and CC2 Results. *J. Chem. Theory Comput.* **2014**, *10*, 3260−3269.

(78) Chakrabarti, S.; Ruud, K. Intermolecular Interaction-Controlled Tuning of the Two-Photon Absorption of Fullerene Bound in a Buckycatcher. *J. Phys. Chem. A* **2009**, *113*, 5485−5488.

(79) Friese, D. H.; Beerepoot, M. T. P.; Ringholm, M.; Ruud, K. Open-Ended Recursive Approach for the Calculation of Multiphoton Absorption Matrix Elements. *J. Chem. Theory Comput.* **2015**, *11*, 1129−1144.

(80) Rudberg, E.; Sałek, P.; Helgaker, T.; Ågren, H. Calculations of two-photon charge-transfer excitations using Coulomb-attenuated density-functional theory. *J. Chem. Phys.* **2005**, *123*, 184108.

(81) Gagliardi, L.; Lindh, R.; Karlström, G. Local properties of quantum chemical systems: the LoProp approach. *J. Chem. Phys.* **2004**, *121*, 4494−4500.

(82) Galván, I. F.; Vacher, M.; Alavi, A.; Angeli, C.; Aquilante, F.; Autschbach, J.; Bao, J. J.; Bokarev, S. I.; Bogdanov, N. A.; Carlson, R. K.; Chibotaru, L. F.; Creutzberg, J.; Dattani, N.; Delcey, M. G.; Dong, S. S.; Dreuw, A.; Freitag, L.; Frutos, L. M.; Gagliardi, L.; Gendron, F.; Giussani, A.; González, L.; Grell, G.; Guo, M.; Hoyer, C. E.; Johansson, M.; Keller, S.; Knecht, S.; Kovačević, G.; Källman, E.; Manni, G. L.; Lundberg, M.; Ma, Y.; Mai, S.; Malhado, J. P.; Malmqvist, P. Å.; Marquetand, P.; Mewes, S. A.; Norell, J.; Olivucci, M.; Oppel, M.; Phung, Q. M.; Pierloot, K.; Plasser, F.; Reiher, M.; Sand, A. M.; Schapiro, I.; Sharma, P.; Stein, C. J.; Sørensen, L. K.; Truhlar, D. G.; Ugandi, M.; Ungur, L.; Valentini, A.; Vancoillie, S.; Veryazov, V.; Weser, O.; Wesołowski, T. A.; Widmark, P. O.; Wouters, S.; Zech, A.; Zobel, J. P.; Lindh, R. OpenMolcas: From Source Code to Insight. *J. Chem. Theory Comput.* **2019**, *15*, 5925−5964.

(83) Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993**, *98*, 5648−5652.

(84) Widmark, P.-O.; Malmqvist, P. k.; Roos, B. O. Density matrix averaged atomic natural orbital (ANO) basis sets for correlated molecular wave functions. I. First row atoms. *Theor. Chim. Acta* **1990**, *77*, 291−306.

(85) Widmark, P.-O.; Persson, B. J.; Roos, B. r. O. Density matrix averaged atomic natural orbital (ANO) basis sets for correlated molecular wave functions. II. Second row atoms. *Theor. Chim. Acta* **1991**, *79*, 419−432.

(86) Pou-Amérigo, R.; Merchán, M.; Nebot-Gil, I.; Widmark, P.-O.; Roos, B. O. Density matrix averaged atomic natural orbital (ANO) basis sets for correlated molecular wave functions. III. First row transition metal atoms. *Theor. Chim. Acta* **1995**, *92*, 149−181.

(87) Zhang, D. W.; Zhang, J. Z. H. Molecular fractionation with conjugate caps for full quantum mechanical calculation of protein-molecule interaction energy. *J. Chem. Phys.* **2003**, *119*, 3599−3605.

(88) Söderhjelm, P.; Ryde, U. How Accurate Can a Force Field Become? A Polarizable Multipole Model Combined with Fragment-wise Quantum-Mechanical Calculations. *J. Phys. Chem. A* **2009**, *113*, 617−627.

(89) Olsen, J. M. H. *PyFraME: Python Framework for Fragment-Based Multiscale Embedding*, version 0.2.0, 2018.

(90) Olsen, J. M.; Aidas, K.; Kongsted, J. Excited States in Solution through Polarizable Embedding. *J. Chem. Theory Comput.* **2010**, *6*, 3721−3734.

(91) List, N. H.; Jensen, H. J. A.; Kongsted, J. Local Electric Fields and Molecular Properties in Heterogeneous Environments through Polarizable Embedding. *Phys. Chem. Chem. Phys.* **2016**, *18*, 10070−10080.

(92) Steindal, A. H.; Beerepoot, M. T. P.; Ringholm, M.; List, N. H.; Ruud, K.; Kongsted, J.; Olsen, J. M. H. Open-ended response theory with polarizable embedding: multiphoton absorption in biomolecular systems. *Phys. Chem. Chem. Phys.* **2016**, *18*, 28339−28352.

(93) Tsuzuki, S.; Honda, K.; Uchimaru, T.; Mikami, M.; Tanabe, K. Origin of Attraction and Directionality of the π/π Interaction: Model Chemistry Calculations of Benzene Dimer Interaction. *J. Am. Chem. Soc.* **2002**, *124*, 104−112.