Computing Excited States of Very Large Systems with Range-Separated Hybrid Functionals and the eXact Integral Simplified Time-Dependent Density Functional Theory (XsTD-DFT)

Marc de Wergifosse*

Theoretical Chemistry Group, Molecular Chemistry, Materials and Catalysis Division (MOST), Institute of Condensed Matter and Nanosciences, Université Catholique de Louvain, Place Louis Pasteur 1, B-1348 Louvain-la-Neuve, Belgium

Received December 4, 2024; E-mail: marc.dewergifosse@uclouvain.be

Abstract: Simplified quantum chemistry (sQC) methods can routinely compute excited states for very large systems in an "all-atom" fashion. They are viable alternative to regular multi-scale schemes. sQC methods have the advantage to account explicitly for all the environment at a QM level. The treatment of charge-transfer states is now improved by the native implementation of range-separated hybrid (RSH) exchange-correlation functionals into the eXact integral simplified time-dependent density functional theory (XsTD-DFT). After benchmarking the RSH XsTD-DFT/TDA scheme, XsTD-DFT(/TDA) UV/vis absorption, circular dichroism (CD), and/or 2PA spectra were directly compared to experiments for four challenging and increasingly-large systems: eYFP model system, Λ-shaped multi-modular D-π-A-π-D'-π-A-π-D chromophore, mixed Donor/Acceptor ligands Pd(II) double cage [$3BF_4$@$Pd_4D_mA_8-m$], and the photoactive yellow protein (PYP). Among results, this study shows that an "all-atom" approach is unavoidable to reproduce absorption and CD spectra of PYP because one of the main transitions involves a local excitation on a tryptophan.

Nowadays, the theoretical characterization of excited states for large systems is mainly performed using multi-scale methods i.e. quantum mechanical/molecular mechanical (QM/MM) schemes.[1,2] However, the interaction of the chromophore with its environment is of QM nature and thus difficult to cope with such treatments.[3] The surroundings could also be involved into the property of interest. Since a decade now, simplified quantum chemistry (sQC) methods[4,5] emerged as an efficient way to compute excited state and response properties of large systems allowing "all-atom" calculations of systems as large as fluorescent proteins[6,7] or fluorescent organic nanoparticles.[8] The simplified time-dependent density functional theory (sTD-DFT)[4,9–14] was originally designed to be used with global hybrid (GH) exchange-correlation functionals (XCFs). Nonetheless, Grimme and coworkers[15] fitted the semi-empirical two-electron integral parameters to be used with range-separated hybrid (RSH) XCFs[16] but TD-DFT equations were still solved for the GH case. Note that the sTD-DFT scheme was only parametrized for few RSH XCFs. Bannwarth and Grimme[11] implemented an alternative to the use of a DFT ground state with the sTD-DFT scheme called the sTD-DFT-xTB method. The ground state is then computed at a semi-empirical level allowing the treatment for systems of few thousands of atoms. Here again, the sTD-DFT step is limited to the GH case. These limitations precluded the application of sQC schemes to properly characterize charge-transfer (CT) states which are of the outmost importance for a plethora of applications.[17–20]

RSH XCFs[16] are an important tool in the quantum chemist's arsenal, especially to characterize CT states. Such functionals improve the treatment of long-range electron-electron interactions for the exchange by incorporating exact Hartree-Fock (HF) exchange using a partitioning scheme.[21] The two-electron operator is separated in short-range (SR) and long-range (LR) parts using the standard error function:[22]

$$\frac{1}{r_{12}} = \underbrace{\frac{1-\mathrm{erf}(\omega r_{12})}{r_{12}}}_{\text{short range}} + \underbrace{\frac{\mathrm{erf}(\omega r_{12})}{r_{12}}}_{\text{long range}}. \quad (1)$$

The LR part of the two-electron operator is used in HF exchange integrals while the SR part for the exchange functional. The mixing parameter ω is adapted for each RSH XCFs. Among RSHs, CAM-B3LYP[23] is one of the most popular functionals that corrects at LR the B3LYP hybrid XC functional. In CAM-B3LYP, a constant amount a_x of HF exchange at SR is also present modifying the splitting of the two-electron operator:

$$\frac{1}{r_{12}} = \underbrace{\frac{1-(a_x+\beta_x\,\mathrm{erf}(\omega r_{12}))}{r_{12}}}_{\text{short range}} + \underbrace{\frac{a_x+\beta_x\,\mathrm{erf}(\omega r_{12})}{r_{12}}}_{\text{long range}}. \quad (2)$$

The amount of LR HF exchange is scaled by a parameter β_x. ωB97X-D3[24] and ωB97M-V[25] functionals share the same formulation. Table 1 provides parameter values for these three RSH XCFs.

Table 1. Parameters for the three RSH XCFs used in this study.

	CAM-B3LYP	ωB97X-D3	ωB97M-V
ω	0.33	0.25	0.3
a_x	0.19	0.1957	0.15
β_x	0.46	0.8043	0.85

To compute excited states using linear response time-dependent functional theory (LR-TD-DFT),[26–32] Casida's equations[33] need to be solved. As we will see in the following, these equations involve the construction of two super-matrices for which orbital energies and XC integral kernel are needed. For lower-rungs functionals, error cancelations often occur between too large orbital energy gaps and too small exchange integral kernel providing rather good excitation energies for valence states but not for Rydberg or CT states.[16] LR-corrected functionals are well known to provide better orbital energies and improve the treatment of exchange terms. In particular, they perform better to describe CT states because they have a rather correct Coulomb asymptotic decay for the exchange potential.[16] Note that

for RSH XCFs that incorporate a fixed amount of exact exchange (e.g. CAM-B3LYP), this decay is correct only for a small part of the exchange potential.

Among RSH XCFs, let's note the existence of short-range corrected SRC1 and SRC2 functionals from Besley et al.[34] that provide an excellent treatment for core-to-valence excitations and will be the subject of a future contribution involving the XsTD-DFT.[5,35] One downside of RSH functionals is that the mixing parameter ω is system-dependent, especially to treat CT states.[36] This parameter is often tuned for specific systems as a function of the ionization potential[37–40] or excitation energies[41] with the inherent limitation of a system specific parametrization.

With the recent introduction of the eXact integral simplified time-dependent density functional theory (XsTD-DFT)[5,35] and the fact that two-electron integrals are no longer evaluated semi-empirically, it is now possible to solve TD-DFT Casida's equations[33] for RSH XCFs allowing a better treatment of CT states by sQC methods. To develop this formalism, I am using the following notations: molecular orbitals (MOs) indices are represented by p, q, r, s, occupied ones by i, j, k, l, unoccupied MOs by a, b, c, d, and atomic orbitals (AOs) by $\alpha, \beta, \gamma, \delta$. The XsTD-DFT[5,35] is developed in the density-matrix-based TD-DFT formalism[26–32] where TD-DFT Casida's equations[33] are simplified as followed:

- Responses of the exchange-correlation functional are neglected.

- Two-electron integrals are evaluated as

$$(ia|jb) \approx \sum_{\alpha\beta} Q_\alpha^{ia} Q_\beta^{jb} \, (\alpha\alpha|\beta\beta) \,, \tag{3}$$

  where Q_α^{ia} transition charges are collected from Löwdin-orthogonalized coefficients $C_{\alpha i}^{low}$ for each basis function α as

$$Q_\alpha^{ia} = \sum_{\alpha} C_{\alpha i}^{low*} C_{\alpha a}^{low}, \tag{4}$$

  and $(\alpha\alpha|\beta\beta)$ are AOs two-electron integrals in the Mulliken notation.

- A single cut-off energy E_{thresh} is used to truncate the space of configuration state functions (CSFs). E_{thresh} represents the spectral range for which excited states need to be computed. Note that sTD-DFT Mataga-Nishimoto-Ohno-Klopman (MNOK) two-electron integrals[42–44] (see Supporting Information (SI)) are still used for the truncation procedure without noticeable changes in results.

In the case of RSH functionals, long-range two electron integrals are approximated as:

$$(ia|\operatorname{erf}(\omega r_{12})|jb) \approx \sum_{\alpha\beta} Q_\alpha^{ia} Q_\beta^{jb} \, (\alpha\alpha|\operatorname{erf}(\omega r_{12})|\beta\beta) \,. \tag{5}$$

The resulting XsTD-DFT equations for RSH XCFs are expressed as:

$$\left[\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B} & \mathbf{A} \end{pmatrix} - \omega \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right] \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} = 0, \tag{6}$$

where $\mathbf{X}$ and $\mathbf{Y}$ are excitation and de-excitation vectors, respectively, and supermatrix elements are defined as

$$\begin{aligned} A_{ia,jb} = \delta_{ij}\delta_{ab}(\epsilon_a - \epsilon_i) + 2(ia|jb)' - a_x(ij|ab)' \\ -\beta_x(ij|\operatorname{erf}(\omega r_{12})|ab)', \end{aligned} \tag{7}$$

and

$$B_{ia,jb} = 2(ia|bj)' - a_x(ib|aj)' - \beta_x(ib|\operatorname{erf}(\omega r_{12})|aj)'. \tag{8}$$

The Tamm-Dancoff approximation[27] neglects the $\mathbf{B}$ supermatrix and gives rise to a simple eigenvalue problem. Note that the RSH XsTD-DFT implementation benefits from the toolbox of linear and quadratic response methods already available in the `stda` program,[45] including UV/vis absorption,[9,10] circular dichroism (CD),[9,10] polarizability,[12] optical rotation,[14] excited state absorption,[13] first hyperpolarizability,[12] and two-photon absorption (2PA).[46] Thanks to its interface with the Libcint library,[47] the XsTD-DFT and XsTDA methods for RSH XCFs were implemented in a development version of the freely-available `stda` program.[45] The interested reader can look to reference 35 to have more details about the XsTD-DFT and XsTDA methods and their implementations.

In this letter, the RSH XsTD-DFT and XsTDA implementations are tested using CAM-B3LYP,[23] ωB97X-D3,[24] and ωB97M-V[25] XCFs. The XsTDA and sTDA excitation energies are benchmarked against TDA results considering the same set of molecules as in reference 35 (Fig.S1 and S2). To assess the performance of the XsTDA scheme to compute intramolecular and intermolecular CT states, TDA, sTDA, and XsTDA excitation energies are compared to reference SCS-CC2/def2-TZVP(-f) values for systems **69** to **77** from the mixed CT test set.[15] The CT excitation energy of system **72** is reported as a function of the distance (R) between both parallel molecular plan of benzene and TCNE for TDA, sTDA, and XsTDA levels of theory and compared to the RI-CC2/aug-cc-pVDZ reference curve.

To illustrate the performance of both RSH XsTDA and XsTD-DFT schemes to reproduce experimental UV/vis absorption, CD, and/or 2PA spectra, two small- to medium-sized systems were selected as well as two challenging large systems (Fig.1). First, the experimental 2PA spectrum of the enhanced yellow protein (eYFP) from reference 48 is compared to results obtained for its chromophore (62 atoms, 268 e^-) at the RI-CC2/aug-cc-pVDZ level from Beerepoot et al.[49] as well as with RI-CAM-B3LYP and XsCAM-B3LYP using the 6-31+g(d) basis set. Then, the experimental 2PA spectrum of a Λ-shaped multi-modular D-π-A-π-D'-π-A-π-D (D=diphenylamine, A=benzothiadizole, and D'=phenothiazine) chromophore (PTZBDT1, 120 atoms, 552 e^-)[50] is compared to the XsωB97X-D3/6-31+G(d) one. In a second step, the experimental UV/vis absorption spectrum of a large mixed D/A ligands Pd(II) double cage $[3BF_4@Pd_4D_mA_{8-m}]$[51] for which the ratio between numbers of donor and acceptor units is unknown is compared to the computed one for the $[3BF_4@Pd_4D_4A_4]^{5+}$ system (395 atoms, 1902 e^-) at the XsCAM-B3LYP/TDA/LAN2DZ/6-31+G(d) level. Finally, experimental UV/vis absorption and CD spectra of the extremely challenging photoactive yellow protein (PYP, 1931 atoms, 7476 e^-)[52] are compared to XsCAM-B3LYP/TDA with the mixed 6-31+G(d)[chromophore + H-bonded residues(see thick sticks, Fig.1)]/6-31G[remaining] basis set and the implicit C-PCM solvent model[53] (water) using the iterative conjugate gradient solver from Lange and Herbert.[54–56] For these calculations, 25854 primitives were used. Computational details are provided in the SI. Note that all 2PA spectra were computed at the XsTD-DFT level while UV/vis absorption and CD ones with the XsTDA method. Both XsTDA and XsTD-DFT methods are efficient enough that all calculations (in-

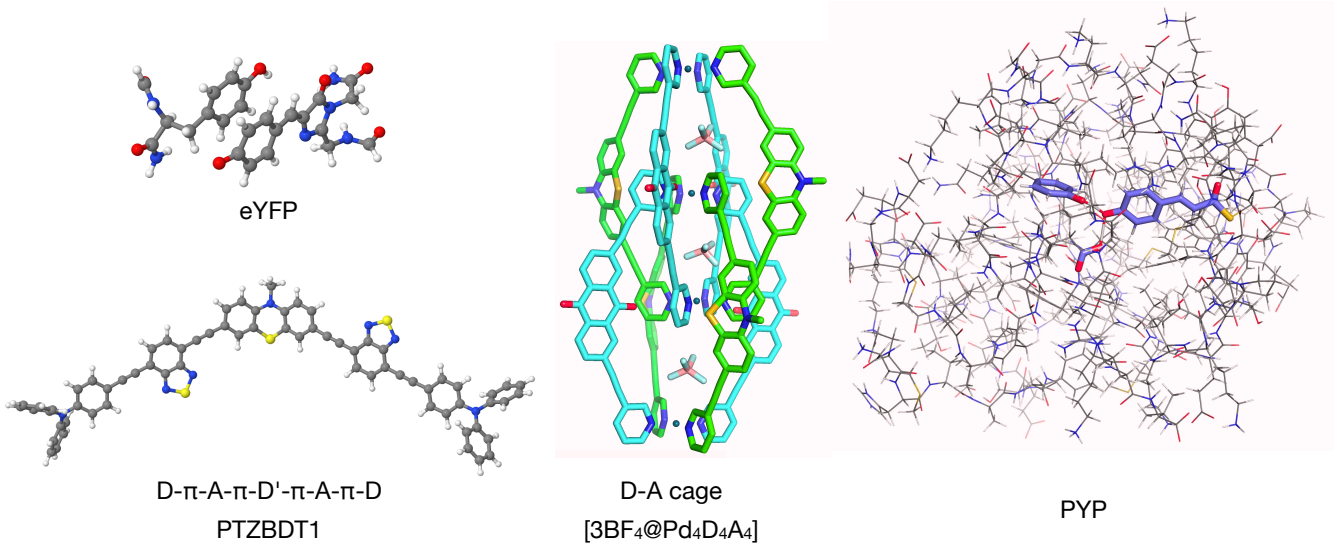


Figure 1. Chemical structures of eYFP chromophore, a multi-modular D-π-A-π-D'-π-A-π-D chromophore (PTZBDT1), a mixed D/A ligands Pd(II) double cage [$3BF_4$@$Pd_4D_4A_4$], and the photoactive yellow protein (PYP).

cluding Q-CHEM[57] SCFs) were performed on a desktop computer (18 CPUs Intel(R) Core(TM) i9-10980XE CPU @ 3.00GHz).

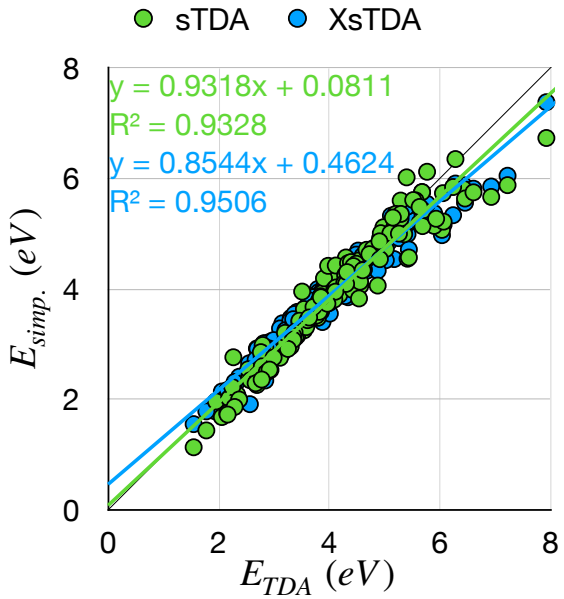


Figure 2. For the two lowest-lying singlet excited states of the set of 77 molecules, correlation graphs for sTDA and XsTDA excitation energies with respect to TDA ones using CAM-B3LYP exchange-correlation functional and the 6-31+G(d) basis set. Linear regression data are also provided.

Let's start discussing results by testing both RSH XsTDA and XsTD-DFT implementations using CAM-B3LYP, ωB97X-D3, and ωB97M-V XCFs for the benchmark set of 77 molecules involving the two first singlet-to-singlet excitations. Figures 2 and S3 present correlation graphs for sTDA and XsTDA excitation energies with respect to TDA ones for the three RSH XCFs considered in this study. Table 2 provides statistical data including mean absolute deviations (MAD), difference between maximum and minimum absolute deviations (AD), and mean deviations (MD). For CAM-B3LYP, on average XsTDA excitation energies deviate less than sTDA ones with respect to TDA values with MADs of 0.22 and 0.27 eV, respectively. The same observation can be done for ωB97X-D3. Note that ωB97M-V has never been parametrized at the sTD-DFT/sTDA levels. This is why no sTDA data are provided for this XCF. For the three functionals, XsTDA MADs with respect to TDA are almost the same, ranking from 0.22 to 0.23 eV. The difference between maximum and minimum AD represent the error spread of the method with respect to the full scheme. Going from sTDA to XsTDA reduces slightly the error spread with respect to TDA. Note that the MDs are negative for all simplified schemes showing on average systematic underestimations of TDA excitation energies.

For the three RSH XCFs, figure S4 presents excitation energies computed at TDA, sTDA, and XsTDA levels of theory for six randomly-chosen compounds from the set that only involve excitations to valence states (two per compound). TDA excitation energies to valence states are robustly reproduced by the XsTDA scheme while they are almost systematically underestimated by the sTDA method. This can also be observed in Fig. 2 and S3. Figure S5 compares TDA, sTDA, and XsTDA excitation energies to reference SCS-CC2 results for the mixed CT set (molecules **69** to **77**)[15] that involves intramolecular (**69** to **71**) and intermolecular (**72** to **77**) CT states. Here again, the XsTDA results are robustly following TDA ones. The sTDA parametrization is excellent for ωB97X-D3 functionals in view of the results while for CAM-B3LYP, TDA excitation energies are systematically underestimated. The comparison between XsωB97M-V and SCS-CC2 excitation energies is striking, especially considering that XsTDA calculations took a maximum of 7 seconds on my desktop computer while the TDA one a bit less than half an hour. For system **72**, figure S6 depicts RI-CC2, TDA, sTDA, and XsTDA excitation energies as a function of the distance (R) between both parallel molecular plan of benzene and TCNE. XsTDA robustly reproduces TDA curves for the three functionals and even outperforms the full scheme at large distance with respect to RI-CC2 results using ωB97X-D3 and ωB97M-V XCFs. These results are extremely encouraging for the treatment of CT states by sQC methods.

After benchmarking both methods, 2PA spectra for the

Table 2. For CAM-B3LYP, ωB97X-D3, and ωB97M-V XCFs, mean absolute deviations (MAD), difference between maximum and minimum absolute deviations (AD), and mean deviations (MD) of sTDA and XsTDA excitation energies with respect to TDA one for the set of 77 molecules.

E (eV)	sCAM-B3LYP	XsCAM-B3LYP	sωB97X-D3	XsωB97X-D3	XsωB97M-V
MAD	0.27	0.22	0.35	0.22	0.23
max AD - min AD	1.35	1.18	1.97	1.23	1.52
MD	-0.20	-0.14	-0.30	-0.12	-0.14

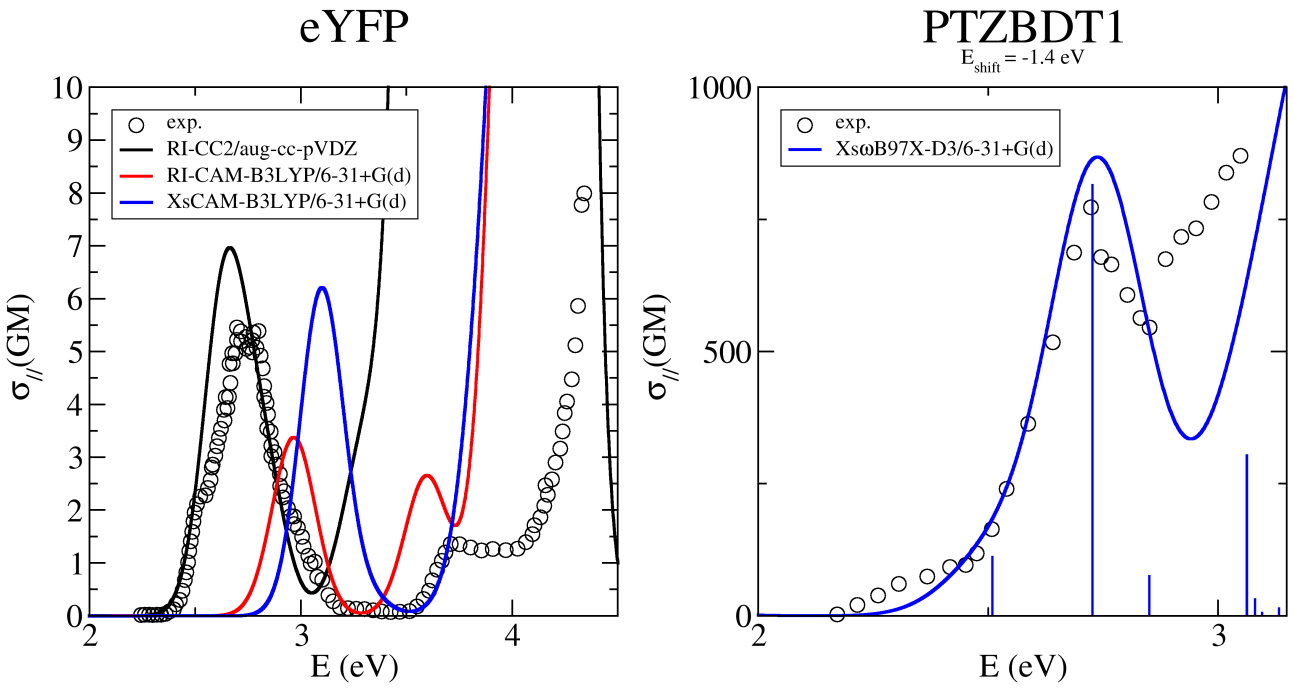


Figure 3. Left panel: experimental 2PA spectrum of eYFP from Drobizhev et al.[48] recorded in water in comparison with computed RI-CC2/aug-cc-pVDZ, RI-CAM-B3LYP/6-31+G(d), and XsCAM-B3LYP/6-31+G(d) spectra for the chromophore in π-staking interaction with a tyrosine. Right panel: experimental 2PA spectrum from Rout et al.[50] recorded in toluene compared the XsωB97X-D3/6-31+G(d) 2PA spectrum shifted by -1.4 eV. A broadening factor of 0.2 eV is employed for convoluted Gaussians.

small- to medium-sized systems were computed at the XsTD-DFT level and directly compared to experiment. In 2014, Beerpoot et al.[49] presented a study about the 2PA of eYFP for which the chromophore is in π-staking interaction with a tyrosine. For this, they used CAM-B3LYP and CC2 and showed that intramolecular CT transitions between the chromophore and the tyrosine are of the outmost importance to interpret the 2PA spectrum. However, the experimental 2PA spectrum from reference 48 only shows transitions below 4.5 eV. These CT transitions occur at energies larger than 4 eV and thus were not measured. Left panel of figure 3 compares the experimental 2PA spectrum to RI-CC2/aug-cc-PVDZ, RI-CAM-B3LYP, and XsCAM-B3LYP ones. No environmental effects or implicit/explicit impact of the protein β-barrel were accounted for the calculations. While shifted by around -0.46 eV with respect to RI-CC2, XsCAM-B3LYP spectrum has a similar shape as the RI-CC2 one. RI-CC2 and XsCAM-B3LYP 2PA cross-sections for the first excited state compare well with experiment. RI-CAM-B3LYP underestimates the experimental 2PA cross-section by almost a factor of 2. Note that the small absorption band around 3.6 eV in the RI-CAM-B3LYP spectrum is also present in both RI-CC2 and XsCAM-B3LYP spectra but hidden by the onset of the larger-energy peak. The comparison with respect to experiment of the XsCAM-B3LYP 2PA spectra is already very good but could be improved by incorporating the protein β-barrel and solvent effects.

The Λ-shaped multi-modular D-π-A-π-D'-π-A-π-D chromophore PTZBDT1 was designed for its particular intramolecular CT character that leads to a large 2PA cross-section in the measurement window.[50] Right panel of figure 3 presents this 2PA spectrum recorded in toluene in comparison with the XsTD-DFT/ωB97X-D3/6-31+G(d) one in gas phase that was shifted by -1.4 eV to match the experimental one. The 2PA cross-section of the main peak is almost perfectly reproduced by the XsTD-DFT calculation with a computed value of 816 GM. Inspecting its natural transition orbitals (Fig.S7) shows that the transition is associated with an intramolecular CT state with a mix between the transition from the hole π orbital of D' to the particle π^* orbital of A and with a lower weight, from the hole π orbital of D to the particle π^* orbital of A. The onset of the peak starting at 2.9 eV is also well-reproduced by the theory. These results could be improved by accounting for solvent effects, reducing the energy shift with respect to experiment but the comparison remains dependent on the choice of the XCFs. Note that I selected ωB97X-D3 XCF for this comparison because the XsCAM-B3LYP 2PA cross-sections are almost two times larger than XsωB97X-D3 ones. Selecting the right XCF for 2PA is still an active subject of discussion in the field.[58–62] On my desktop computer, the 2PA XsTD-DFT calculation took only 3.33 minutes for 120 atoms, thanks to the ultra-fast 2PA implementation at the simplified level.[46]

Finally, for the two large and challenging systems of this study, UV-vis and/or CD spectra were evaluated at the XsTDA level of theory and compared to experimental results. Clever and coworkers[51] designed a supramolecular structure composed of two interpenetrated cages. Each cage involves two Pd(II) atoms linked together by donor (phenothiazine) or acceptor (anthraquinone) ligands. They were able to synthesize homo-octameric double cages as well as mixed-ligand double cages $[3BF_4@Pd_4D_mA_{8-m}]^{5+}$ (Fig.1) for which they were not able to control the ratio between the number of donor and acceptor groups. Figure 4 presents

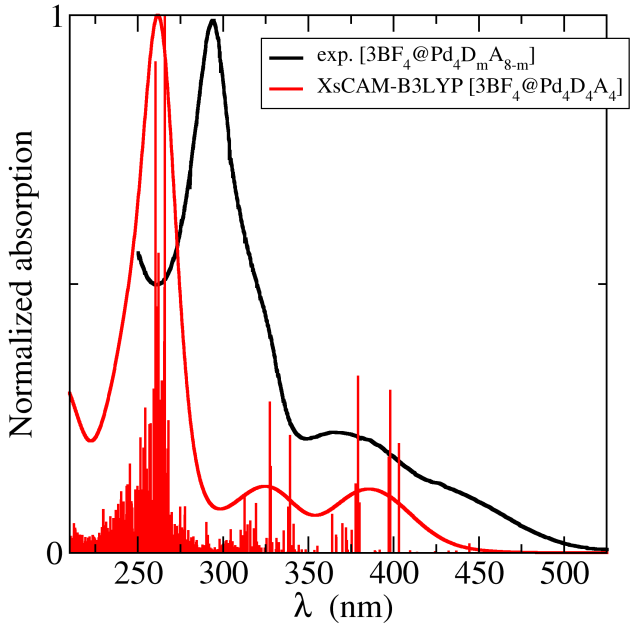


Figure 4. Experimental UV-vis spectrum from Clever and coworkers[51] recorded in acetonitrile compared the XsCAM-B3LYP/TDA/LAN2DZ/6-31+G(d) spectrum. A broadening factor of 0.2 eV was used for convoluted Gaussians.

the experimental UV-vis spectrum for this extremely challenging system recorded in acetonitrile. In this spectrum, three main absorption bands are present. From measurements done for the homo-octameric double cages with only acceptor or donor ligands, they identify that the first band involves donor ligands while the second band acceptor ligands. The largest absorption band at 295 nm according to them is dominated by intramolecular $\pi \rightarrow \pi^*$ transitions. To model this system, I used a well-defined double cages $[3BF_4@Pd_4D_4A_4]^{5+}$ with two acceptor and two donor ligands per cage. Excited states were computed at the XsCAM-B3LYP/TDA/LAN2DZ/6-31+G(d) level in gas phase. Figure 4 shows the convoluted spectrum. The agreement with experiment is salient while the main peak is only shifted by 30 nm because of missing effects such as the impact of the solvent, considering only singly-excited configurations,... The spectral shape looks falsely simple because each computed band is composed of a large number of excitations (see oscillator strengths in Fig.4). Computed states 18, 51, and 223 correspond to the three band maxima, respectively. To confirm or infirm assignments from Clever and coworkers,[51] I computed NTOs (Fig.S8) at the same level of theory. All transitions clearly show CT character. The NTO pair associated with the transition from the ground state to the 18th excited state clearly shows a CT transition from a hole π orbital on a phenothiazine to a particle π^* orbital on a pyridine from another nearby donor ligand of the other cage. For the second absorption band, we have two main NTO pairs: one associated with a CT transition from a hole π orbital on a phenothiazine to a particle π^* orbital located on another acceptor ligand linked to the same Pd atom and the other one for a local $\pi \rightarrow \pi^*$ transition involving two nearby anthraquinone. The transition to state 223 is dominated by a transition from a π orbital located on a phenothiazine to a d orbital from a Pd linked to this ligand. For this case, the XsTDA calculation was able to reproduce the experiment while providing unambiguous assignments for main absorp-

tion bands which are not fully confirming ones from Clever and coworkers.[51]

For the largest structure of this study PYP, because the system contains 1931 atoms and 7476 e^-, mixed-basis sets were employed to be still able to perform DFT ground state calculations with Q-CHEM[57] on my desktop computer (see Computational Details in SI). The right panel of figure S9 displays the UV-vis absorption spectra computed at the XsCAM-B3LYP/TDA/C-PCM(water) level of theory with 3-21G, 6-31+G(d)/3-21G, 6-31G, and 6-31+G(d)/6-31G basis sets including 17709, 18107, 25635, and 25854 primitives, respectively. Clearly, one can observe that diffuse and polarization sets for the chromophore play a more important role than simply going from 3-21G to 6-31G with respect to the 6-31+G(d)/6-31G result. Excitation energies to the first excited state and corresponding oscillator strengths are already well described using small basis sets. These interesting basis set effects might advocate for the use of a mixed basis set approach for large systems with central chromophores. CAM-B3LYP was selected because results compare better to experiment than ωB97M-V and PBE0 (see Fig.S9, left panel). Figure 5 compares experimental UV/vis absorption and CD spectra of PYP recorded by Borucki et al.[52] in water to XsCAM-B3LYP/TDA/6-31+G(d)/6-31G//C-PCM(water) ones. A sytematic energy shift of -1.5 eV was applied to match experiment. The agreement with respect to experiment is striking, especially to reproduce spectral shapes and relative peak energy positions. Only the rotatory strength for the first excited state is underestimated with respect to experiment. As shown in figure S10 for which TDA and XsTDA UV-vis spectra were computed for the chromophore and both H-bonded residues only, it is necessary to account for the full environment to be able to describe the second peak correctly. NTOs for transitions from the ground state to the 1st, 21st, and 23rd excited states are displayed in figure S11. Table S1 provides excitation energies, oscillator strengths, rotatory strengths, and wavelengths shifted by -1.5 eV for the 1st, 21st, 23rd, and 34th excited states of PYP at the XsCAM-B3LYP/TDA/6-31+G(d)/6-31G//C-PCM(water) level of theory. The transition to the first excited state appears in both UV/vis absorption and CD spectra is a $\pi \rightarrow \pi^*$ transition located on the chromophore with some CT character. The second main peak in the UV/vis absorption spectrum is due to a transition to state 23rd which is also a $\pi \rightarrow \pi^*$ transition but from a tryptophan residue in the protein. This directly explains why computations on the chromophore only were not able to describe this peak correctly. Only an "all-atom" approach as the XsTDA one could describe such transitions. Indeed, using multi-scale schemes, it is very likely that such residues will only be accounted for at the MM level. This result advocates for the use of "all-atom" schemes such as the XsTDA one to correctly describe fluorescent proteins for which absorbing residues are also present and important to take into account. In the CD spectrum, the 21st state is responsible for the positive peak located at 286 nm. It involves a $\pi \rightarrow \pi^*$ transition located on the tyrosine H-bonded to the chromophore. The time determining step for the XsTDA is the DFT ground state computed with Q-CHEM.[57] The CAM-B3LYP/6-31+G(d)/6-31G//C-PCM(water) ground state calculation took less than 38 hours while the XsTDA calculation with $E_{thresh} = 6.5$ eV took 51 minutes.

In summary, I provided in this study the first native implementations of RSH XCFs with sQC methods at both XsTDA and XsTD-DFT levels of theory. After benchmarking the RSH XsTDA method for a set of 77 molecules and showing the excellent performance of the XsTDA scheme at robustly reproducing TDA excitation energies for CT states, XsTDA and XsTD-DFT methods were used to model UV/vis absorption, CD, and/or 2PA spectra of four increasingly large systems for which experimental data are available. Excellent agreements with respect to experiment were obtained in the limit of the usual performance of RSH XCFs. Meaning that as expected from such functionals large energy shifts ($\leq$ -1.5 eV) were applied to reproduce experimental spectra, i.e. for PTZBDT1 and PYP. Nonetheless, experimental spectral shapes were very well reproduced by the theory for the four systems. For eYFP, the XsTD-DFT 2PA spectra for the chromophore and the stacking tyrosine gives a similar comparison with respect to experiment as when employing the RI-CC2 method but with an energy shift of -0.46 eV and at a tiny fraction of the computational cost. For the Λ-shaped multi-modular D-π-A-π-D'-π-A-π-D chromophore, the experimental 2PA cross-section for the large CT excitation responsible of the main absorption band was almost perfectly reproduced by the XsTD-DFT calculation. The CT character of the state was confirmed by the analysis of the NTOs. Regarding the challenging mixed-ligand double cages $[3BF_4@Pd_4D_mA_{8-m}]^{5+}$ involving four Pd(II) atoms, here again the XsTDA scheme provides an excellent comparison to experiment for the UV-vis spectrum as well as unambiguous assignments for the three main absorption bands. For the most absorbing band, the NTO analysis confirms a $\pi \rightarrow d$ CT state from a donor phenothiazine ligand to a linked Pd, in contradiction with the intramolecular $\pi \rightarrow \pi^*$ transitions suggested by Clever and coworkers.[51] Finally, the most challenging case of this study was to reproduce at an "all-atom" sQC level both experimental UV-vis absorption and CD spectra for PYP, a fluorescent protein composed of 1931 atoms. Both spectral shapes were well reproduced at the XsCAM-B3LYP/TDA level. Only the rotatory strength for the transition from the ground state to the first excited state was underestimated by the theory with respect to experiment. More surprisingly, the NTO analysis showed that only an "all-atom" calculation could have reproduced those spectra because one of the main transitions is dominated by a local $\pi \rightarrow \pi^*$ located on a tryptophan from the protein. All these results with respect to experiment strongly advocate for the use of both RSH XsTDA and XsTD-DFT methods as new tools in the quantum chemist's arsenal to characterize optical properties of large systems in an "all-atom" fashion.

Supporting Information Available

See Supporting Information for: a) Theory Supplement, b) Computational Details, c) Chemical Structures of Systems **1** to **77**, d) Correlation Graphs, e) Excitation Energy plots for Selected Valence and CT States, f) Natural Transition Orbitals for PTZBDT1, $[3BF_4@Pd_4D_mA_{8-m}]^{5+}$, and PYP, g) Extra Computed UV-Vis Spectra for PYP

Acknowledgement The author thanks his PhD student, Marilù G. Maraldi who performed extra RI-TD-DFT calculations on the eYFP chromophore and Prof. Benoît Champagne for his valuable comments on the manuscript.

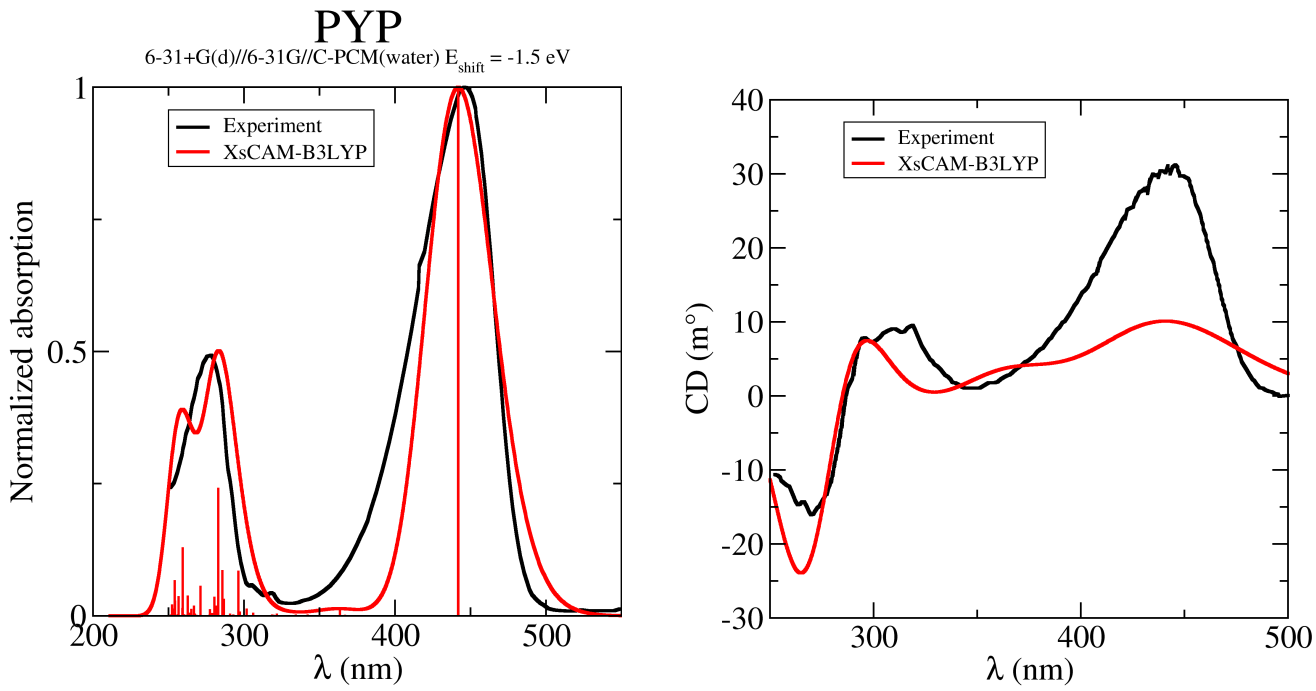


Figure 5. UV/vis absorption and CD spectra of PYP recorded by Borucki et al.[52] in water in comparison with XsCAM-B3LYP/TDA/6-31+G(d)/6-31G//C-PCM(water) ones. A broadening factor of 0.2 eV was used for convoluted Gaussians for the UV/vis absorption spectrum while 0.3 eV was chosen for the CD spectrum. CD data are reported as ellipticity in millidegrees.

References

(1) Zuehlsdorff, T. J.; Isborn, C. M. Modeling Absorption Spectra of Molecules in Solution. *Int. J. Quant. Chem.* **2019**, *119*, e25719.

(2) Giovannini, T.; Cappelli, C. Continuum Vs. Atomistic Approaches to Computational Spectroscopy of Solvated Systems. *Chem. Comm.* **2023**, *59*, 5644–5660.

(3) Giovannini, T.; Egidi, F.; Cappelli, C. Molecular Spectroscopy of Aqueous Solutions: A Theoretical Perspective. *Chem. Soc. Rev.* **2020**, *49*, 5664–5677.

(4) de Wergifosse, M.; Grimme, S. Perspective on Simplified Quantum Chemistry Methods for Excited States and Response Properties. *J. Phys. Chem. A* **2021**, *125*, 3841–3851.

(5) Löffelsender, S.; Beaujean, P.; de Wergifosse, M. Simplified Quantum Chemistry Methods to Evaluate Non-Linear Optical Properties of Large Systems. *WIREs Comp. Mol. Sci.* **2023**, *14*, e1695.

(6) Seibert, J.; Bannwarth, C.; Grimme, S. Biomolecular Structure Information from High-Speed Quantum Mechanical Electronic Spectra Calculation. *J. Am. Chem. Soc.* **2017**, *139*, 11682–11685.

(7) Beaujean, P.; Champagne, B.; Grimme, S.; de Wergifosse, M. All-Atom Quantum Mechanical Calculation of the Second-Harmonic Generation of Fluorescent Proteins. *J. Phys. Chem. Lett.* **2021**, *12*, 9684–9690.

(8) Lescos, L.; Beaujean, P.; Tonnelé, C.; Aurel, P.; Blanchard-Desce, M.; Rodriguez, V.; de Wergifosse, M.; Champagne, B.; Muccioli, L.; Castet, F. Self-Assembling, Structure and Nonlinear Optical Properties of Fluorescent Organic Nanoparticles in Water. *Phys. Chem. Chem. Phys.* **2021**, *23*, 23643–23654.

(9) Grimme, S. A Simplified Tamm-Dancoff Density Functional Approach for the Electronic Excitation Spectra of Very Large Molecules. *J. Chem. Phys.* **2013**, *138*, 244104.

(10) Bannwarth, C.; Grimme, S. A Simplified Time-Dependent Density Functional Theory Approach for Electronic Ultraviolet and Circular Dichroism Spectra of Very Large Molecules. *Comput. Theor. Chem.* **2014**, *1040-1041*, 45 – 53.

(11) Grimme, S.; Bannwarth, C. Ultra-Fast Computation of Electronic Spectra for Large Systems by Tight-Binding Based Simplified Tamm-Dancoff Approximation (sTDA-XTB). *J. Chem. Phys.* **2016**, *145*, 054103.

(12) de Wergifosse, M.; Grimme, S. Nonlinear-Response Properties in a Simplified Time-Dependent Density Functional Theory (sTD-DFT) Framework: Evaluation of the First Hyperpolarizability. *J. Chem. Phys.* **2018**, *149*, 024108.

(13) de Wergifosse, M.; Grimme, S. Nonlinear-Response Properties in a Simplified Time-Dependent Density Functional Theory (sTD-DFT) Framework: Evaluation of Excited-State Absorption Spectra. *J. Chem. Phys.* **2019**, *150*, 094112.

(14) de Wergifosse, M.; Seibert, J.; Grimme, S. Simplified Time-Dependent Density Functional Theory (sTD-DFT) for Molecular Optical Rotation. *J. Chem. Phys.* **2020**, *153*, 084116.

(15) Risthaus, T.; Hansen, A.; Grimme, S. Excited States Using the Simplified Tamm-Dancoff-Approach for Range-Separated Hybrid Density Functionals: Development and Application. *Phys. Chem. Chem. Phys.* **2014**, *16*, 14408–14419.

(16) Tsuneda, T.; Hirao, K. Long-Range Correction for Density Functional Theory. *WIREs Comp. Mol. Sci.* **2014**, *4*, 375–390.

(17) Chen, X.-K.; Coropceanu, V.; Brédas, J.-L. Assessing the Nature of the Charge-Transfer Electronic States in Organic Solar Cells. *Nat. Comm.* **2018**, *9*, 5295.

(18) Coropceanu, V.; Chen, X.-K.; Wang, T.; Zheng, Z.; Brédas, J.-L. Charge-Transfer Electronic States in Organic Solar Cells. *Nat. Rev. Mater.* **2019**, *4*, 689–707.

(19) Imahori, H.; Kobori, Y.; Kaji, H. Manipulation of Charge-Transfer States by Molecular Design: Perspective from "Dynamic Exciton". *Acc. Mater. Res.* **2021**, *2*, 501–514.

(20) Baharfar, M.; Hillier, A. C.; Mao, G. Charge-Transfer Complexes: Fundamentals and Advances in Catalysis, Sensing, and Optoelectronic Applications. *Adv. Mater.* **2024**, *36*, 2406083.

(21) Iikura, H.; Tsuneda, T.; Yanai, T.; Hirao, K. A Long-Range Correction Scheme for Generalized-Gradient-Approximation Exchange Functionals. *J. Chem. Phys.* **2001**, *115*, 3540–3544.

(22) Savin, A. In *Recent Developments and Applications of Modern Density Functional Theory*; Seminario, J., Ed.; Theoretical and Computational Chemistry; Elsevier, 1996; Vol. 4; pp 327–357.

(23) 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.

(24) Lin, Y.-S.; Li, G.-D.; Mao, S.-P.; Chai, J.-D. Long-Range Corrected Hybrid Density Functionals with Improved Dispersion Corrections. *J. Chem. Theory Comput.* **2013**, *9*, 263–272.

(25) Mardirossian, N.; Head-Gordon, M. ωB97M-V: A Combinatorially Optimized, Range-Separated Hybrid, Meta-GGa Density Functional with VV10 Nonlocal Correlation. *J. Chem. Phys.* **2016**, *144*, 214110.

(26) Zangwill, A.; Soven, P. Density-Functional Approach to Local-Field Effects in Finite Systems: Photoabsorption in the Rare Gases. *Phys. Rev. A* **1980**, *21*, 1561–1572.

(27) Hirata, S.; Head-Gordon, M. Time-Dependent Density Functional Theory Within the Tamm-Dancoff Approximation. *Chem. Phys. Lett.* **1999**, *314*, 291 – 299.

(28) Furche, F. On the Density Matrix Based Approach to Time-Dependent Density Functional Response Theory. *J. Chem. Phys.* **2001**, *114*, 5982–5992.

(29) Wang, F.; Yam, C. Y.; Chen, G. Time-Dependent Density-Functional Theory/localized Density Matrix Method for Dynamic Hyperpolarizability. *J. Chem. Phys.* **2007**, *126*, 244102.

(30) Thorvaldsen, A. J.; Ruud, K.; Kristensen, K.; Jørgensen, P.; Coriani, S. A Density Matrix-Based Quasienergy Formulation of the Kohn-Sham Density Functional Response Theory Using Perturbation- and Time-Dependent Basis Sets. *J. Chem. Phys.* **2008**, *129*, 214108.

(31) Zahariev, F.; Gordon, M. S. Nonlinear Response Time-Dependent Density Functional Theory Combined with the Effective Fragment Potential Method. *J. Chem. Phys.* **2014**, *140*, 18A523.

(32) Parker, S. M.; Rappoport, D.; Furche, F. Quadratic Response Properties from TDDFT: Trials and Tribulations. *J. Chem. Theory Comput.* **2018**, *14*, 807–819.

(33) Casida, M. E. *Time-Dependent Density Functional Response Theory for Molecules*; 1995; pp 155–192.

(34) Besley, N. A.; Peach, M. J. G.; Tozer, D. J. Time-Dependent Density Functional Theory Calculations of Near-Edge X-Ray Absorption Fine Structure with Short-Range Corrected Functionals. *Phys. Chem. Chem. Phys.* **2009**, *11*, 10350–10358.

(35) de Wergifosse, M.; Grimme, S. The eXact Integral Simplified Time-Dependent Density Functional Theory (XsTD-DFT). *J. Chem. Phys.* **2024**, *160*, 204110.

(36) Körzdörfer, T.; Sears, J. S.; Sutton, C.; Brédas, J.-L. Long-Range Corrected Hybrid Functionals for π-Conjugated Systems: Dependence of the Range-Separation Parameter on Conjugation Length. *J. Chem. Phys.* **2011**, *135*, 204107.

(37) Baer, R.; Livshits, E.; Salzner, U. Tuned Range-Separated Hybrids in Density Functional Theory. *Annu. Rev. Phys. Chem.* **2010**, *61*, 85–109.

(38) Karolewski, A.; Kronik, L.; Kümmel, S. Using Optimally Tuned Range Separated Hybrid Functionals in Ground-State Calculations: Consequences and Caveats. *J. Chem. Phys.* **2013**, *138*, 204115.

(39) Imamura, Y.; Kobayashi, R.; Nakai, H. Linearity Condition for Orbital Energies in Density Functional Theory (IV): Determination of Range-Determining Parameter. *Int. J. Quant. Chem.* **2013**, *113*, 245–251.

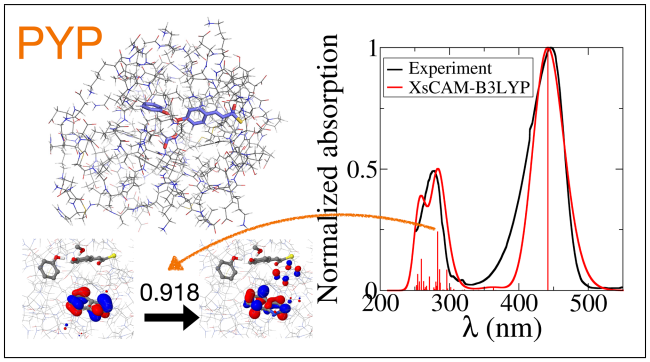
(40) Jacquemin, D.; Moore, B. I.; Planchat, A.; Adamo, C.; Autschbach, J. Performance of an Optimally Tuned Range-Separated Hybrid Functional for 0–0 Electronic Excitation Energies. *J. Chem. Theory Comput.* **2014**, *10*, 1677–1685.

(41) Lin, Z.; Van Voorhis, T. Triplet Tuning: A Novel Family of Non-Empirical Exchange–Correlation Functionals. *J. Chem. Theory Comput.* **2019**, *15*, 1226–1241.

(42) Nishimoto, K.; Mataga, N. Electronic Structure and Spectra of

Some Nitrogen Heterocycles. *Z. Phys. Chem.* **1957**, *12*, 335–338.
(43) Ohno, K. Some Remarks on the Pariser-Parr-Pople Method. *Theor. Chim. Acta* **1964**, *2*, 219–227.
(44) Klopman, G. A Semiempirical Treatment of Molecular Structures. II. Molecular Terms and Application to Diatomic Molecules. *J. Am. Chem. Soc.* **1964**, *86*, 4550–4557.
(45) The stda code can be obtained from: https://github.com/grimme-Lab/stda/, date of access 9/2/2024. `https://github.com/grimme-lab/stda/`.
(46) de Wergifosse, M.; Beaujean, P.; Grimme, S. Ultrafast Evaluation of Two-Photon Absorption with Simplified Time-Dependent Density Functional Theory. *J. Phys. Chem. A* **2022**, *126*, 7534–7547.
(47) Sun, Q. Libcint: An Efficient General Integral Library for Gaussian Basis Functions. *J. Comp. Chem.* **2015**, *36*, 1664–1671.
(48) Drobizhev, M.; Makarov, N. S.; Tillo, S. E.; Hughes, T. E.; Rebane, A. Two-Photon Absorption Properties of Fluorescent Proteins. *Nat. Meth.* **2011**, *8*, 393–399.
(49) 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.
(50) Rout, Y.; Cesaretti, A.; Ferraguzzi, E.; Carlotti, B.; Misra, R. Multiple Intramolecular Charge Transfers in Multimodular Donor–Acceptor Chromophores with Large Two-Photon Absorption. *J. Phys. Chem. C* **2020**, *124*, 24631–24643.
(51) Frank, M.; Ahrens, J.; Bejenke, I.; Krick, M.; Schwarzer, D.; Clever, G. H. Light-Induced Charge Separation in Densely Packed Donor–Acceptor Coordination Cages. *J. Am. Chem. Soc.* **2016**, *138*, 8279–8287.
(52) Borucki, B.; Otto, H.; Meyer, T. E.; Cusanovich, M. A.; Heyn, M. P. Sensitive Circular Dichroism Marker for the Chromophore Environment of Photoactive Yellow Protein: Assignment of the 307 and 318 Nm Bands to the N $\rightarrow \pi^*$ Transition of the Carbonyl. *J. Phys. Chem. B* **2005**, *109*, 629–633.
(53) Barone, V.; Cossi, M. Quantum Calculation of Molecular Energies and Energy Gradients in Solution by a Conductor Solvent Model. *J. Phys. Chem. A* **1998**, *102*, 1995–2001.
(54) Lange, A. W.; Herbert, J. M. a Smooth, Nonsingular, and Faithful Discretization Scheme for Polarizable Continuum Models: The Switching/Gaussian Approach. *J. Chem. Phys.* **2010**, *133*, 244111.
(55) Lange, A. W.; Herbert, J. M. Symmetric Versus Asymmetric Discretization of the Integral Equations in Polarizable Continuum Solvation Models. *Chem. Phys. Lett.* **2011**, *509*, 77–87.
(56) Herbert, J. M.; Lange, A. W.; Cui, Q.; Ren, P.; Meuwly, M. Polarizable Continuum Models for (bio) Molecular Electrostatics: Basic Theory and Recent Developments for Macromolecules and Simulations. *Many-Body Effects and Electrostatics in Biomolecules* **2016**, 363–416.
(57) Epifanovsky, E. et al. Software for the Frontiers of Quantum Chemistry: An Overview of Developments in the Q-Chem 5 Package. *J. Chem. Phys.* **2021**, *155*, 084801.
(58) Chołuj, M.; Alam, M. M.; Beerepoot, M. T. P.; Sitkiewicz, S. P.; Matito, E.; Ruud, K.; Zaleśny, R. Choosing Bad Versus Worse: Predictions of Two-Photon-Absorption Strengths Based on Popular Density Functional Approximations. *J. Chem. Theory Comput.* **2022**, *18*, 1046–1060.
(59) Grotjahn, R.; Furche, F. Gauge-Invariant Excited-State Linear and Quadratic Response Properties Within the Meta-Generalized Gradient Approximation. *J. Chem. Theory Comput.* **2023**, *19*, 4897–4911.
(60) Ahmadzadeh, K.; Li, X.; Rinkevicius, Z.; Norman, P.; Zaleśny, R. Toward Accurate Two-Photon Absorption Spectrum Simulations: Exploring the Landscape Beyond the Generalized Gradient Approximation. *J. Phys. Chem. Lett.* **2024**, *15*, 969–974.
(61) Naim, C.; Zaleśny, R.; Jacquemin, D. Two-Photon Absorption Strengths of Small Molecules: Reference CC3 Values and Benchmarks. *J. Chem. Theory Comput.* **2014**, *20*, 9093–9106.
(62) Grotjahn, R.; Furche, F. Comment On: "Toward Accurate Two-Photon Absorption Spectrum Simulations: Exploring the Landscape Beyond the Generalized Gradient Approximation". *J. Phys. Chem. Lett.* **2024**, *15*, 6237–6240.

TOC Graphic


PYP
0.918
Normalized absorption
1
0.5
0
200
300
400
500
λ (nm)
Experiment
XsCAM-B3LYP