ChemPhysChem

Research Article
doi.org/10.1002/cphc.202401121

Chemistry Europe
European Chemical Societies Publishing

www.chemphyschem.org

All-Atom Quantum Mechanical Methodologies for One- and Two-Photon Absorption of Realistic Systems

Sarah Löffelsender, Marilù G. Maraldi, and Marc de Wergifosse*

All-atom quantum mechanics (AQM) methodologies are assessed to evaluate one- and two-photon absorption (1PA and 2PA) of realistic systems. All-atom single structure QM (ASQM) and dynamic structure QM (ADQM) methodologies are discussed. These workflows are possible thanks to developments in simplified quantum chemistry methods and in particular with both sTD-DFT-xTB and dt-sTD-DFT-xTB schemes. The ASQM scheme is tested to compute the 1- and 2PA of two proteins: bacteriorhodopsin and iLOV. Results show that the ASQM methodology is able to describe higher-energy transitions involving π-conjugated amino acids such as tryptophan or tyrosine. Then, two variants of the ADQM workflow are evaluated to reproduce the 1- and 2PA of the flavin mononucleotide in aqueous solution, involving either Boltzmann ensemble of conformers in implicit solvent (ADQM-Boltz.) or snapshots of molecular dynamics of explicitly solvated systems (ADQM-MD). Spectra computed with the ADQM-MD approach provide striking comparisons with respect to experiment, while the ADQM-Boltz. approach provides little change with respect to the ASQM workflow.

1. Introduction

The direct environment of a molecule modifies its ground and excited state properties including geometry, energy, dynamics, and reactivity.[1] For example, the theoretical description of solvatochromism should involve a realistic representation of surrounding molecules around the solute to faithfully reproduce one- and two-photon absorption (1PA and 2PA) experimental spectra.[1,2] For intermolecular interactions such as hydrogen bonds, π-stacking, or halogen bonds, an explicit treatment of the environment should naturally be included into the methodology.[2] The computational modeling of highly flexible systems, for example, molecules in solution, necessitates the integration of dynamic structural effects.[1,2] Such treatment can be computationally prohibitive for large molecules or molecular assemblies.[3] **Figure 1** shows two currently established workflows[1,2] to compute excited states of molecules in solution that include dynamic structural effects. In the first workflow (left side), the impact of the environment is modeled implicitly using a solvent model applied to i) an isolated molecule or ii) a cluster composed of the chromophore and few explicit solvent molecules that includes microsolvation interactions while keeping a moderate computational cost. Dynamic structural effects are captured implicitly by sampling the conformational space, and Boltzmann averaged spectra are computed. The second workflow (right side of Figure 1) accounts explicitly for direct surroundings of the chromophore. Molecular dynamics (MD) simulations are performed to compute solute–solvent (chromophore–environment) configurations and account for temperature effects.[1,2,4,5] These simulations are usually conducted using classical, Monte–Carlo, or quantum mechanical/molecular mechanical (QM/MM) MD simulations.[2,6] For each snapshot, excited states are typically computed using a QM/MM framework (reducing the number of explicitly treated atoms drastically). An averaged spectra incorporating each snapshot is then obtained. Note that to account for polarization effects and to converge excited state properties,[7] some authors suggested to include $>$250 atoms in the QM region[8] or at least two solvation shells.[1] This places a significant computational demand on the excited state calculation. In both workflows, two main steps can be distinguished: i) sampling geometries and ii) characterizing excited states with the environment accounted for either implicitly or explicitly.

In 1976, Warshel and Levitt[9] introduced the QM/MM method to treat large molecular systems in a consistent way. For the development of such multiscale models, they won the Nobel Prize in Chemistry in 2013 with Karplus. In the QM/MM approach, the system considered responsible for the target property is computed quantum mechanically while surroundings are treated classically. These schemes enable the treatment of explicitly solvated systems as well as chromophores embedded in more complex matrices such as proteins. Nowadays, a plethora of schemes exists for the MM part,[2,6,10–22] involving mechanical embedding (ME), electrostatic embedding (EE), or polarizable embedding (PE). They rely on a classical description of the solute–solvent (or chromophore–environment) electrostatic interactions, either neglecting (ME, EE) or accounting for (PE) important mutual polarizations between the solute and the environment. Nevertheless, solute–solvent (or chromophore–environment)

S. Löffelsender
Mulliken Center for Theoretical Chemistry
¬Clausius Institut für Physikalische und Theoretische Chemie
University of Bonn
Beringstr. 4, 53115 Bonn, Germany

M. G. Maraldi, M. de Wergifosse
Theoretical Chemistry Group, MolecularChemistry
¬Materials and Catalysis Division (MOST)
¬Institute of Condensed Matter and Nanosciences
Université Catholique de Louvain
B-1348 Louvain-la-Neuve, Belgium
E-mail: marc.dewergifosse@uclouvain.be

Supporting information for this article is available on the WWW under https://doi.org/10.1002/cphc.202401121

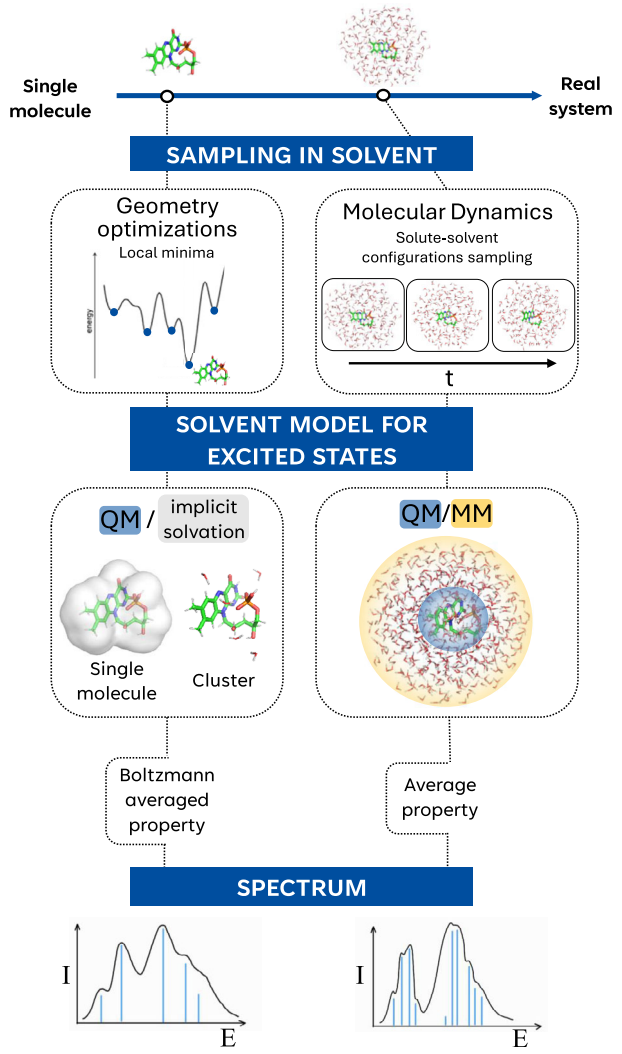

Figure 1. Two current workflows to compute excited state spectra in solution or complex environments that account for dynamic structural effects.

interactions have a quantum nature, which are particularly difficult to describe with QM/MM models.[6] Higher-level approaches also exist such as the QM/effective fragment potential (EFP)[23–27] or quantum embedding schemes.[28–30]

2PA is a nonlinear optical (NLO) effect that allows to use lower energy light sources with respect to 1PA. Both incident photons have a lower energy than the excited state transition energy (half of it, if degenerate). It exhibits an improved spatial resolution and an increased penetration depth (up to 1 nm, compared to 50–80 μm in normal confocal laser scanning microscopy) as well as reduced photobleaching. These characteristics motivated the use of 2PA as a noninvasive imaging technique for numerous medical fields, including cancer[31–33] or kidney researches,[34–36] neuroscience,[37] and even to probe neurological functions of (fixed or freely moving) living animals.[38–42] 2PA is also used in data storage devices[43,44] and information processing.[45] These areas of research could benefit from the design of new systems with enhanced 2PA cross sections (σ^{2PA}).

Calculating 2PA spectra is a challenging task as it formally involves to evaluate a third-order molecular response property.[46,47] Computationally efficient expressions to evaluate σ^{2PA} can be derived from residue analysis of the cubic response function.[46,48] For the degenerate case, 2PA strengths can also be extracted from the residue of the quadratic response function.[46] Since the seminal work of Olsen and Jørgensen[46] in 1985, hierarchy of methods were implemented along the years including couple cluster (CC) and density functional theory (DFT) methods. Nanda and Krylov[49] implemented one of the most formally accurate scheme available with the equation of motion coupled cluster singles and doubles (EOM-CCSD) method. Hattig et al.[48,50,51] implemented CCS, CCSD, CC3, CC2 and RI-CC2 methods to evaluate 2PA. Cost-efficient (RI-)CC2 schemes are often used as a reference to evaluate 2PA strengths.[4,52–59] CC methods are still computationally demanding and are only available to treat rather small systems (<100 atoms). Alternatively, time-dependent DFT methods[60–62] were implemented for medium-sized systems.

QM/MM schemes were often used to evaluate 2PA of chromophores in solution or complex matrices. Model systems of explicitly solvated chromophores were employed to characterize the effect of the solvent on 2PA spectral features,[63–67] and two-photon solvatochromism,[68] as well as to assess dynamic structural effects.[69] Some of them provided computational results directly comparable to experiments.[65,66,68] For example, Olesiak–Banska et al.[65] studied the 2PA of the Hoechst chromophore in water solution showing the emergence of a low-energy band in the 2PA spectra due to dimerization of the dye in solution. The QM/MM family of methods was also used to evaluate the 2PA of chromophores in protein environments.[70–72] For channel rhodopsin mutants[73] and green fluorescent proteins,[72] design guidelines based on QM/MM studies were proposed to enhance their σ^{2PA}. Murugan and Zaleśny[71] used a CC2/MM approach to model the 2PA properties of Parkinson's diagnostic probes targeting monoamine oxidase B. To improve the modeling of theoretical 2PA spectra, Kongsted and coworkers[74] showed for DNA intercalators the importance to include non-electrostatic repulsion into the embedding scheme, reducing the electron density leakage from the QM to the MM region. Olsen et al.[75] showed that very accurate explicitly polarizable embedding potentials for proteins can be efficiently designed using system fragmentation strategies for QM/MM molecular response calculations. Excited state properties can also be highly sensitive to chromophore geometries.[2,76–79] For example, Palczewska et al.[79] examined the impact of distorting bovine rhodopsin on 2PA using QM/MM (QM, chromophore; MM, residue charges within a 5 Å distance). They demonstrated that the deformation of the chromophore as well as surrounding atoms and their charges have significant impacts on the 2PA spectrum, complementing the study of Valsson et al.[8] on 1PA.

As already stressed above, the nature of interactions between the chromophore and its direct environment has a quantum nature. De facto, such interactions are particularly difficult to be described with QM/MM models especially if important long-range interactions or direct H-bonds between layers occur.[6] A significant limitation associated with such schemes is to interface both QM and MM regions, although several protocols exist.[13] A viable alternative to QM/MM schemes for the

computation of excited states is to directly treat most of the system in a QM fashion while remaining parts are accounted for by a solvent model. This is the so-called "brute force approach" as termed by Giovannini et al.[2] for which we prefer the terminology "all-atom QM (AQM) methodology." To compute excited states and response properties of large systems, an AQM methodology containing explicitly up to several thousands of atoms is now possible thanks to recent developments in the field of simplified quantum chemistry methods.[80–91] The realm of simplified methods was first introduced by Grimme[80] in 2013 with the simplified time-dependent DFT (sTD-DFT) variant using the Tamm–Dancoff approximation (sTDA). Shortly after, this was extended to the sTD-DFT method by Bannwarth and Grimme.[81] To compute excited states of ultra large systems (up to 5000 atoms), an approach called sTD-DFT-xTB combines an extended tight binding (xTB) ground state[82] with the sTD-DFT method. The computation of response properties with the sTD-DFT scheme was introduced by de Wergifosse et al. including the polarizability,[83] optical rotation,[86] excited state absorption,[84] first hyperpolarizability,[83] and the ultra-fast evaluation of 2PA.[90] Among applications, Seibert et al.[87] assessed dynamic structural effects on the first hyperpolarizability of very flexible tryptophan-rich peptides as well as the gramicidin A; Beaujean et al.[88] proposed an AQM methodology to compute the first hyperpolarizabilty of two fluorescent proteins (FPs) iLOV and the bacteriorhodopsin (bR), showing striking agreement with respect to experiment. The dual-threshold sTD-DFT (dt-sTD-DFT) was introduced to reduce the computational cost of nonlinear response calculations by truncating in a different manner the space of singly excited configurations for the chromophore and the remaining of the protein.[88] Last year, one of us with the help of the eXact integral sTD-DFT (XsTD-DFT)[91,92] using the CAM-B3LYP range-separated hybrid functional showed that to reproduce experimental 1PA and circular dichroism spectra of the photoactive yellow protein (PYP), an AQM methodology is absolutely necessary because one of the main peak is due to a local $\pi \rightarrow \pi^*$ transition on a tryptophan and thus not involving the chromophore.[93]

The bacteriorhodopsin (3835 atoms) is a transmembrane protein of Halobacterium salinarum.[94–97] It is a light-driven proton pump that facilitates the transfer of protons from the external to the internal cellular environment. Prior studies have demonstrated that the chromophore, a retinal, interacts considerably with nearby residues (Trp86 and Tyr185) upon excitation,[98–100] resulting in a remarkable high molar extinction coefficient.[101] Moreover, reports indicate an exceptionally high 2PA cross section,[102] with possible applications in data storage.[43,44] bR undergoes photobleaching upon 2P excitation, due to the accessibility of photoproducts beyond the classical photocycle.[103,104] Note that the retinal chromophore from bR is a different molecule with respect to the well-studied retinal vitamin A from the human eye.

iLOV is an improved variant of the light, oxygen, or voltage domain of the plant blue light receptor.[105] This small FP (1849 atoms) originally engineered to monitor virus infection and spread owes its photophysical properties to its noncovalently bound chromophore flavin mononucleotide (FMN). Theoretical and experimental studies[4,76–78,106–113] on FMN established that its surroundings influence significantly its excited state properties. **Figure 2** presents experimental 1PA spectra of FMN in different environments (water, methanol, and two similar proteins iLOV and miniSOG), showing the sensitivity of FMN to its environment. At a first glance, trends among different types of surroundings can be rationalized in terms of the polarity of the environment. For 1PA spectra, flavins were found to be sensitive to the planarity of the isoalloxazine ring.[76,111,113] Already subtle distortions, that is, the well-known "butterfly" bending of the ring system, causes substantial changes in the spectra,[76,110,111] for example, shifting of the excitation energy. This was attributed to the mixing of $n \rightarrow \pi^*$ and $\pi \rightarrow \pi^*$ transitions.[76,110,113] In the low-energy region (between 450 and 350 nm), FMN has dark singlet and triplet $n \rightarrow \pi^*$ states especially sensitive to the environment.[77,110,111,113] Upon significant distortions due to surrounding atoms, those forbidden transitions are then allowed, modifying the spectra. Depending on the environment, the structure of flavin isoalloxazine rings can be planar[76] or distorted.[110] Kar et al.[78] suggested the existence of a wide distribution of angles in flavins encapsulated in protein environments. In a theoretical 1PA and 2PA investigation, List et al.[4] used a QM/MM scheme with PE and accounted for explicit dynamic structural effects to examine the behavior of FMN in aqueous solution and in the miniSOG protein. Comparisons with respect to experimental 1PA and 2PA spectra indicate that spectral shapes are rather well reproduced, while peak energy positions are blue-shifted.

The objective of this study is to establish possible AQM workflows to compute 1- and 2PA of molecules in solution and complex environments (see **Figure 3**). Note that in this contribution, the terminology "realistic systems" is used to illustrate that not only calculations on single molecules or small systems are preformed but also including explicitly large parts of the environment as well as dynamic structural effects (when possible) to provide model systems closer to reality. The general idea is to rely only on QM levels of theory and using a similar approach as some of us did for the evaluation of the first hyperpolarizabilty of FPs.[88] Figure 3 gives a general overview of the methodology we will apply in this work. A first protocol is proposed for FPs

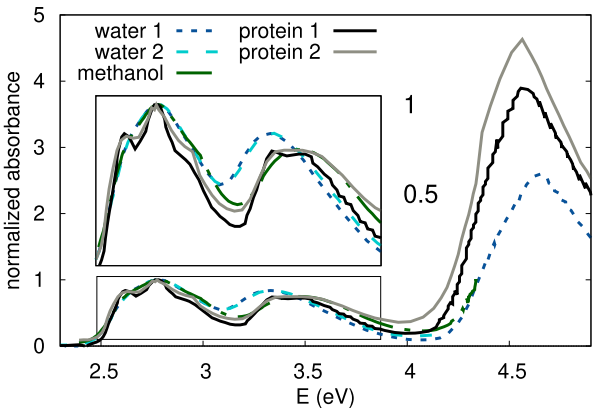

Figure 2. Experimental 1PA spectra of the FMN in different surroundings (water 1,[106] water 2,[4] methanol,[107] protein 1 (iLOV),[108] and protein 2 (miniSOG)[109]). All spectra were normalized to one for the lowest-energy band around 2.79 nm.

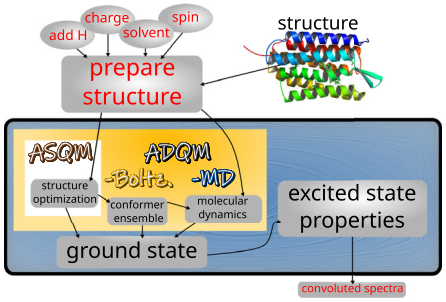

Figure 3. General workflow to obtain excited state spectra. Red font marks critical user-defined steps. In this work, each step inside the black framed box considers all atoms quantum mechanically.

where we use a single structure approach, neglecting dynamic structural effects: an all-atom single structure QM (ASQM) methodology. For this, we used the same systems as in Beaujean et al.[88] that is, bR (≈3850 atoms) and iLOV (≈2000 atoms) (see **Figure 4** left and middle) optimized at the ONIOM ωB97X-D/6-31G*:GFN2-xTB/GBSA(water) level of theory. The sTD-DFT-xTB method[82] is selected to compute excited states as well as 1PA and 2PA spectra directly considering entire structures. For both steps, the generalized born/surface area (GBSA) model[114–116] is used. The challenging systems to assess the ASQM approach (bR and iLOV) were selected because experimental 1PA and 2PA spectra are available in the literature.[102,108,109,117–119] With these systems, we are at the edge of what is computationally feasible nowadays for such scheme.

A second protocol is proposed for more structurally dynamic systems such as chromophores in solution. This protocol is named all-atom dynamic structure QM (ADQM). In that case, two approaches were tested to account for dynamic structural effects adapting usually admitted workflows (Figure 1) to AQM procedures. First, solvent effects are treated implicitly and the conformational space is screened at the tight-binding GFN2-xTB level using the conformer/rotamer sampling tool CREST.[120] Afterward, the GFN2-xTB ensemble is refined with the r^{2}SCAN-3c[121–123] DFT method using CENSO.[124,125] Boltzmann averaged 1- and 2PA spectra are determined using the sTD-DFT-xTB/GBSA method, accounting implicitly for dynamic structural effects. This variant of the ADQM protocol is named ADQM-Boltz. Second, solvent effects are incorporated by a sphere of explicit solvent molecules around the chromophore. Dynamic structural effects are considered by running mixed classical/QM GFN2-xTB/GBSA MD simulations. This approach computes forces analytically at the QM GFN2-xTB level of theory to solve Newton's equation of motions. Thus, only trajectories are computed classically.[126] For each uncorrelated snapshot selected from these MD simulations, 1- and 2PA spectra are computed at the sTD-DFT-xTB/GBSA level of theory and averaged. This variant of the ADQM protocol is called ADQM-MD. Note that in our MD simulations, nuclei move classically which might impact high-frequency motions and narrowing spectral absorption bands. Strategies to modify spectral width or asymmetry of states exists,[127–129] that is, using the nuclear-ensemble approximation that uses a stochastic ensemble of normal coordinates generated by a Wigner distributions. This approach seems limited to rather small systems. Lukeš et al.[130] showed that increasing the MD temperature could be used as a parameter to provide better peak broadenings equivalent to those produce by a Wigner sampling. A feature that will be investigated in future studies as a refinement of the present ADQM protocol. FMN (Figure 4 right) was selected for this first exploratory study to assess the proposed AQM protocol that includes dynamical effects of the solvent and the chromophore. Results are compared to available experimental 1PA and 2PA spectra for FMN in aqueous solution.[4,105–108]

Note that while AQM protocols are developed to better capture the complexity of the experiment by providing the most realistic system we can model nowadays, comparing results from such calculations to experiment is still a tremendous task. Our calculations which neglect vibronic effects provide excitation energies that might need to be shifted and strengths that need to be convoluted by arbitrary functions (either Lorentzian or Gaussian) with an arbitrary width and converted into macroscopic cross sections. Converting 2PA strengths to macroscopic cross sections depend on experimental setups as stressed by Beerepoot et al.[54] and this information is not always available in the literature. Experimental error bars are not always provided neither. On the top of this, 2PA cross sections are dependent on the square of the excitation energy. Thus, what we compare to experiment is not always completely well-defined and clear.

This study was also the opportunity to test the dt-sTD-DFT-xTB method to compute excited states, which has never been

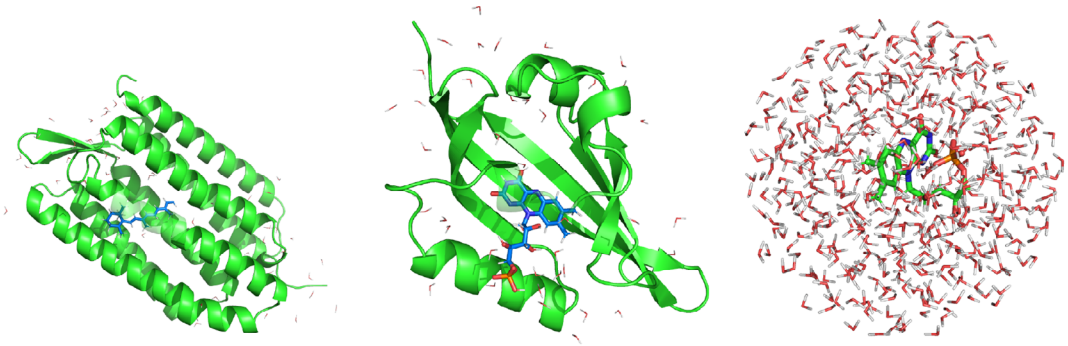

Figure 4. Chemical sketches of bacteriorhodopsin (left panel), iLOV (middle panel), and the flavin mononucleotide surrounded by a shell of 12 Å of explicit water molecules (right panel).

done until now. This scheme seems particularly relevant to reduce the computational cost when a large number of excited states calculations need to be performed, for example, when considering hundreds snapshots from MD runs. We will demonstrate its advantage and illustrate how it can reduce the computational effort while maintaining accuracy in comparison to regular sTD-DFT-xTB calculations.

This study is organized as follows: we first give a concise overview of the theory behind the dt-sTD-DFT as well as how to compute 1- and 2PA at the sTD-DFT level. Then, the computational details employed throughout this work are given (Section 3). In Section 4, results are discussed. Finally, the conclusions and outlooks are given in Section 5.

2. Theory

The dt-sTD-DFT method was developed to reduce the computational cost of sTD-DFT calculations to evaluate the first hyperpolarizability of FPs.[88,91] Here, we extend the reach of this scheme to determine excited states and to evaluate 2PA cross sections. This section briefly recalls the theory behind the sTD-DFT method to compute excited states[80,81] and 2PA cross sections.[90] Details are also given about the dt-sTD-DFT scheme.[88,91] In this study, we will use $i,j,\ldots$ for occupied, $a,b,\ldots$ for virtual and $p,q,\ldots$ for molecular orbitals (MO) of any kind.

The sTD-DFT method is based on Casida's equations:

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B} & \mathbf{A} \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} = \begin{pmatrix} \omega & 0 \\ 0 & -\omega \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} \tag{1}$$

where **A** and **B** are orbital rotation Hessian matrices, **X** and **Y** are the excitation and de-excitation vectors, and ω represents the eigenvalue diagonal matrix with the dimension of the number of roots (or states) and describes the excitation energies. In 2013, Grimme[80] proposed three simplifications to construct **A** and **B** efficiently: i) the exchange-correlation kernel f_{XC} in **A** and **B** supermatrices is neglected to avoid time-intensive numerical integration; ii) the two-electron integrals are approximated by short-ranged damped Coulomb interactions of Löwdin transition charges; and iii) the single excitation space is truncated, leading to a significant speed-up in time with only a minor loss of accuracy.[80,81] Matrix elements of **A**′ and **B**′ in sTD-DFT read the following.

$$\mathbf{A}'_{ia,jb} = \delta_{ij}\delta_{ab}(\varepsilon_a - \varepsilon_i) + \sum_{A,B}^{N} (2q^A_{ia}q^B_{jb}\Gamma^K_{AB} - q^A_{ij}q^B_{ab}\Gamma^J_{AB}) \tag{2}$$

$$\mathbf{B}'_{ia,jb} = \sum_{A,B}^{N} (2q^A_{ia}q^B_{bj}\Gamma^K_{AB} - a_x q^A_{ib}q^B_{aj}\Gamma^K_{AB}) \tag{3}$$

where two-electron MO integrals written in the chemists' notations $(pq|rs) = \iint \Psi^*_p(\boldsymbol{r}_1)\Psi_q(\boldsymbol{r}_1)\frac{1}{r_{12}}\Psi^*_r(\boldsymbol{r}_2)\Psi_s(\boldsymbol{r}_2)\mathrm{d}\boldsymbol{r}_1\mathrm{d}\boldsymbol{r}_2$ are approximated as follows

$$(pq|rs) \approx \sum_{A}^{N}\sum_{B}^{N} q^A_{pq}q^B_{rs}\Gamma_{AB} \tag{4}$$

Transition charges q^A_{pq} are obtained from a Löwdin population analysis,[131] and the Γ_{AB} function damps their interaction at short distance using the Mataga–Nishimoto–Ohno–Klopman formula.[132–134] For a Coulomb-type integral $(ij|ab)'$, the damping formula takes the form

$$\Gamma^J_{AB} = \left(\frac{1}{(R_{AB})^{y_J} + (a_x\eta)^{-y_J}}\right)^{\frac{1}{y_J}} \quad \text{with} \quad \eta = \frac{\eta(A)+\eta(B)}{2} \tag{5}$$

where R_{AB} denotes the distance between atom A and B, y_J is a linear function depending on the amount of Fock exchange a_x, and η is the average chemical hardness of two atoms. Exchange type integrals $(ib|aj)'$ are damped using a slightly different equation.

$$\Gamma^K_{AB} = \left(\frac{1}{(R_{AB})^{y_K} + \eta^{-y_K}}\right)^{\frac{1}{y_K}} \tag{6}$$

where y_K is again a linear function that depends on the amount of Fock exchange. Both Γ_{AB} functions were determined as $y_J = 0.20 + 1.83a_x$ and $y_K = 1.42 + 0.48a_x$, respectively.[80,81]

To truncate the configurations interaction space (related to the size of **A**′ and **B**′), one single energy threshold E_{thr} is used as parameter. E_{thr} represents the maximum excitation energy to be computed to represent the 1PA spectral range. It was shown that a careful truncation of the CI space decreases computational costs in time and memory effectively without loosing much accuracy.[135–138] As a preliminary step, the MO active space is truncated as follows.

$$\begin{aligned} \varepsilon_{min} &= \varepsilon_{LUMO} - 2(1 + 0.8a_x)E_{thr} \\ \varepsilon_{max} &= \varepsilon_{HOMO} + 2(1 + 0.8a_x)E_{thr} \end{aligned} \tag{7}$$

where ε_{HOMO} and ε_{LUMO} are energies of the highest occupied (HO) and lowest unoccupied (LU) molecular orbitals. Primary configuration state functions (P-CSFs), describing the single excitations from occupied MO i to unoccupied MO a, are selected when

$$A'_{ia,ia} \leq E_{thr} \tag{8}$$

For remaining CSFs with $A'_{jb,jb} > E_{thr}$, using a perturbative approach, secondary CSFs (S-CSFs) are only considered if they are coupled significantly to the P-CSFs space

$$E^{(2)}_{jb} = \sum_{ia}^{P-CSFs} \frac{|A'_{ia,jb}|^2}{A'_{jb,jb} - A'_{ia,ia}} > 10^{-4}E_h \tag{9}$$

Consequently, the total number of CSFs is the sum of P- and S-CSFs selected. For FMN at the sTD-DFT-xTB level, the space of 12 096 CSFs is drastically reduced to 342 CSFs using a E_{thr} value of 7 eV and 1413 CSFs for $E_{thr} = 10$ eV.

When considering larger systems with thousands of atoms, computational costs increase drastically. The dt-sTD-DFT scheme gives a solution to truncate even further the space of CSFs for large systems with a central chromophore.[88,91] The dual threshold approach divides the system into two subsystems, which are treated by two different energy thresholds. The motivation is

simple: Considering a chromophore inside its explicit environment, for example, solvent, the response of the chromophore (few atoms) is much more significant than the response of the surroundings (many atoms). Therefore, CSFs involving the chromophore will impact much more the excited state manifold. Thus, a larger energy threshold (E_{high}) should be employed for the chromophore, while the surroundings are covered with a lower energy threshold (E_{low}), reducing computational costs. Note that E_{high} should at least represent the spectral range.

In the dt-sTD-DFT procedure, the truncation of the MO space is done as before (see Equation (7)) considering E_{high} as the energy threshold. Then, the occupied MO space is split into two parts, the high level (e.g., the chromophore) and the low level ones (e.g., the solvent). If an occupied MO has a density ζ higher than 10% on atoms from the high level part, it is included into the high-level occupied MO space. The remaining occupied MOs are treated within the low-level occupied MO space. For a MO i, we have the following.

$$\zeta_i = \sum_{\alpha \in \text{high level part}} C_{\alpha i}^2 \tag{10}$$

$$\zeta_i > 0.1 \rightarrow E_{\text{high}} \tag{11}$$

$$\zeta_i \leq 0.1 \rightarrow E_{\text{low}} \tag{12}$$

$C_{\alpha i}$ are LCAO coefficients and ζ_i is the electron density of the MO i on the high level part of the system.

P-CSFs are selected using Equation (8) with E_{high} and E_{low}, for single excitations, respectively, involving the high level and low level occupied MO spaces. This leads to the P-CSFs-H space if $A_{i_{\text{high}}a, i_{\text{high}}a} \leq E_{\text{high}}$ and to P-CSFs-L space if $A_{i_{\text{low}}a, i_{\text{low}}a} \leq E_{\text{low}}$. To prune the remaining space, only $j_{\text{low}} \rightarrow b$ CSFs with

$$E_{\text{low}} < A'_{j_{\text{low}}b, j_{\text{low}}b} < 2(1 + 0.8a_x)E_{\text{low}} \tag{13}$$

are considered for the low level part, while all remaining $j_{\text{high}} \rightarrow b$ CSFs are selected for the high level part. S-CSFs are then selected if

$$E_{jb}^{(2)} = \sum_{ia}^{P-\text{CSF}s-(H+L)} \frac{|A'_{ia,jb}|^2}{A'_{ia,ia} - A'_{jb,jb}} \tag{14}$$

Finally, the total amount of CSFs is the sum of all primary CSFs from the high- and low-level parts as well as the S-CSFs, that is, $\#\text{CSF}s = \#P-\text{CSF}s - H + \#P-\text{CSF}s - L + \#S-\text{CSF}s$.

To compute a 1PA spectrum at the (dt-)sTD-DFT level, the transition dipole moment $\vec{\mu}_{0\nu}^L$ between the ground state and the excited state ν is evaluated using $\mathbf{X}'$ and $\mathbf{Y}'$. Considering the restricted case in the length formalism, the transition dipole moment reads

$$\overrightarrow{\mu}_{0\nu}^L = \sqrt{2}\sum_{ia} \overrightarrow{\mu}_{ia}(X'^{\nu}_{ia} + Y'^{\nu}_{ia}) \tag{15}$$

where $\vec{\mu}_{ia} = \langle \psi_i | \vec{r} | \psi_a \rangle$. The associated oscillator strength is obtained as follows.

$$f_{0\nu} = \frac{2}{3}\omega_{\nu 0}\vec{\mu}_{0\nu} \cdot \vec{\mu}_{\nu 0} \tag{16}$$

Oscillator strengths are then convoluted with Gaussian functions with a sole damping factor Γ at half width at $1/e$ maximum for all transitions to obtain the 1PA spectrum.

A 2PA spectrum is computed with the (dt-)sTD-DFT method by first evaluating two-photon (2P-) transition dipole moments $M_{\zeta\eta}^{0\rightarrow n}$ using the following expression.[90]

$$M_{\zeta\eta}^{0\rightarrow n} = -A + B \tag{17}$$

where A and B read

$$A = \sum_{\text{perm.}\xi,\zeta,\eta} \left\{ \sum_{aij} X'_{n,ai}[\mu_{\zeta,ij}(1 - \delta_{\zeta\eta})] Y'_{\eta,aj}(-\omega_n/2) \right\} \tag{18}$$

and

$$B = \sum_{\text{perm.}\xi,\zeta,\eta} \left\{ \sum_{iab} X'_{n,ai}[\mu_{\zeta,ab}(1 - \delta_{\zeta\eta})] Y'_{\eta,bi}(-\omega_n/2) \right\} \tag{19}$$

$\mathbf{X}'_\eta(-\omega_n/2)$ and $\mathbf{Y}'_\eta(-\omega_n/2)$ are linear response vectors obtained at the frequency $-\omega_n/2$ from extra linear response calculations for each transition, and perm.ξ,ζ,η stands for permutations between cartesian coordinates and their respective frequencies. From transition dipole moments, the rotationally averaged 2PA transition strength $\delta^{2\text{PA}}$[139,140] is evaluated as follows.

$$\langle \delta^{2\text{PA}} \rangle = \frac{F}{30}\sum_{\zeta,\eta} S_{\zeta\zeta,\eta\eta} + \frac{G}{30}\sum_{\zeta,\eta} S_{\zeta\eta,\zeta\eta} + \frac{H}{30}\sum_{\zeta,\eta} S_{\zeta\eta,\eta\zeta} \quad \text{with} \quad S_{\zeta\eta,\xi\nu} = M_{\zeta\eta}^{0\rightarrow n} M_{\xi\nu}^{0\rightarrow n} \tag{20}$$

where F, G, and H are parameters which depend on the experimental setup. In the case of parallel linearly polarized incident light, F, G, and H equal 2. All 2PA spectra provided in this contribution consider this case. To compute 2PA spectra, macroscopic 2PA cross-section σ^{2PA} are computed as follows.

$$\sigma^{2PA} = \frac{N\pi^3 \alpha a_0^5 (2\omega)^2}{c} \langle \delta^{2PA} \rangle G(2\omega, \omega_n, \Gamma) \tag{21}$$

where the parameter N depends on the experimental setup to compare with (single or double beam experiment, $N = 1$ or 2), α depicts the fine structure constant, a_0 denotes the Bohr radius, c is the speed of light, and G represents a line-shape function. Gaussians are usually used for Γ to model 2PA spectra in solution.[54]

$$G(2\omega, \omega_n, \Gamma) = \frac{\sqrt{\ln 2}}{\sqrt{\pi}\Gamma} \exp\left[-\ln 2 \left(\frac{2\omega - \omega_n}{\Gamma} \right)^2 \right] \tag{22}$$

where Γ is the half-width at half-maximum (HWHM). Note that while an identical Γ factor is usually used for all transitions,[49,52,90,141–144] it was shown that 2P-transitions have often different broadenings,[145,146] originating from different interactions. In the following, broadening factors are given separately for each system and are chosen to best reproduce experiment. Note that for average ADQM-MD spectra, smaller damping factors are used to capture spectral features from out of equilibrium structures that

are naturally broadening absorption bands. All along the article, 2PA spectra are presented as function of 2ω.

3. Computational Details

To assess the ASQM protocol, structures of bR (bacteriorhodospin) and iLOV were taken from Beaujean et al.[88] They were optimized at the ONIOM ωB97X-D[147,148]/6-31G*[149–154]:GFN2-xTB[155]/GBSA(water)[114–116] level of theory and include few explicit water molecules (57 in bR and 46 in iLOV) around the protein. The ωB97X-D range-separated hybrid exchange-correlation functional was used at the DFT level for the ONIOM high layer including the chromophore and the surrounding amino acids and water molecules in a 4 Å radius. The remaining of the protein including external water molecules was accounted for in the ONIOM low layer with the GFN2-xTB method.[155] Excited states as well as δ^{2PA} were computed at the (dt-)sTD-DFT-xTB level of theory.[82] Note that the implicit GBSA(water) model was used to further account for solvent effects for ground state xTB calculations. For each sTD-DFT-xTB 2PA calculation, δ^{2PA} were computed for the 20 first excited states and converted to σ^{2PA} using Equation (21). Note that computing σ^{2PA} for the 20 first excited states is enough to match the experimental 2PA spectral range for all systems considered in this study.

The ADQM protocol includes dynamic structural effects considering its two variants: ADQM-Boltz. and ADQM-MD. In the ADQM-Boltz. methodology for water solution, the deprotonated FMN conformer/rotamer ensemble (CREs) was determined at the GFN2-xTB level of theory using CREST 2.12.[120] Solvent effects were accounted for using the GBSA implicit solvent model. The CRE for FMN in vacuum was computed in the same manner, excluding any solvent model. Both CREs were refined at the DFT level: first, by using the B97-D3[156,157]/def2-SV(P)[158] scheme for a cheap prescreening, and second the composite r^2SCAN-3c[121–123] method to reoptimize geometries. Implicit solvent effects were included using the DCOSMO-RS(water)[159,160] model. Final Gibbs energy rankings are obtained using the r^2SCAN-3c electronic energy + G_{mRRHO}[161–163](GFN2-xTB)//r^2SCAN-3c for the ensemble in gas phase as well as the r^2SCAN-3c electronic energy + COSMO-RS[H_2O][164,165] + G_{mRRHO}(GFN2[GBSA]-bhess)//r^2SCAN-3c[DCOSMO-RS][159,160] in implicit water. This procedure is performed with the command line energetic sorting algorithm CENSO 1.2.0,[124,125] using Turbomole V7.6,[166–168] the xtb program package V6.6.1,[125] and the COSMOtherm 16 program package(G_{solv} option, BP_TZVP_C30_1601 parametrization, $T = 298.15$ K, $p = 1$ atm).[169]

Alternatively, dynamic structural effects were accounted for by taking snapshots from MD simulations, giving the ADQM-MD protocol. Accounting explicitly for solvent effects, water molecules were added to FMN lowest-energy conformer using spheres of increasingly large radii of 7, 9, 10, 11, and 12 Å with the program Chimera[170] ("shell" solvation and TIP3PBOX solvation model). Each structure were preoptimized at the GFN2-xTB/GBSA(water) level to remove possible tensions due to the solvation process, giving better starting points for MD simulations. 1 ns long GFN2-xTB/GBSA(water) MDs were run with a time step of 4 fs to determine an NVT ensemble at 298.15 K. Note that the SHAKE algorithm was used to constrain all bonds, and we applied four times the mass of hydrogen to all hydrogen atoms. Two polynomial wall potentials were set as follows: 1) One spherical potential to constrain the whole system and prevent water molecules to "escape," mimicking the "glass" of a real world experiment. Spheres of radius values of 11.6, (22.0), 13.2 (25.0), 14.7 (27.7), 15.3 (29.0), and 16.6 (31.45) Å (Bohr) were chosen for the potential when considering 7, 9, 10, 11, and 12 Å of water shells, respectively. Note that in the following, we always refer to FMN solvated in *X* Å water shell as "*X* Å." 2) A second wall potential was set around a central atom of FMN with a sphere radius of 2.6 Å (5.0 Bohr) to keep the referential of the molecule centered and avoid drifting apart.

A longer MD simulation (3 ns) was also conducted for the 10 Å water shell to assess the impact of the MD time length on spectra. To select uncorrelated snapshots from MD simulations, all trajectories were evaluated using the autocorrelation function (ACF) of the program TRAVIS.[171,172] Additionally, the block averaging method was used to get the smallest time step between two uncorrelated configurations. This procedure led to 118 (7 Å), 98 (9 Å), 237 (10 Å), 134 (11 Å), and 166 (12 Å) snapshots, respectively. To complement discussions about FMN in gas phase, a 11 ns long GFN2-xTB MD simulation on FMN lowest-energy conformer was run using the same conditions as described above. 1521 snapshots were taken.

For each structure generated with the second protocol, xTB/GBSA(water) ground states for sTD-DFT-xTB calculations were computed. Excitation energies, oscillator strengths, and δ^{2PA} were determined at the (dt-)sTD-DFT-xTB level of theory. For each conformer or MD snapshot, 1PA and 2PA spectra were generated by convoluting oscillator strengths or δ^{2PA} with a Gaussian lineshape. Smaller damping factors were employed for spectra of MD snapshots before taking the average to resolve structural influences (see Section 4).

Furthermore, RI-CC2[173]/aug-cc-pVDZ[174] calculations were performed to compute 1- and 2PA spectra for retinal and FMN lowest-energy conformer obtained with r^2SCAN-3c in gas phase and implicit solvent (FMN only). Turbomole 7.7.1[175] was used to compute eight RI-CC2 transition energies and cross sections using default parameters. Note that we used a value of 1000 for the maximum dimension of the reduced space of the CC linear equations (maxred). A convergence threshold of 10^{-6} is used for the numerical Laplace transformation to compute δ^{2PA}.

The xtb program package[125] version 6.5.1 was used for all GFN2-xTB calculations. All-atom xTB ground states were calculated with the xtb4stda program.[176] (dt-)sTD-DFT-xTB calculations were performed using a development version of the std2 package (previously branded stda).[80,90,177] Note that all along this study, vibronic effects are not accounted for because the gradient is not implemented yet in std2.[80,90,177]

4. Results and Discussions

The main goal of this work is to assess two AQM protocols considering the effect of the environment explicitly without and with dynamic structural effects: ASQM and ADQM, respectively. This section is divided into two parts. Using the ASQM protocol, we

first discuss 1- and 2PA of two very large systems: bR and iLOV. This first part is also the opportunity to discuss the dt-sTD-DFT approach to compute excited states for such systems. Second, we investigate the influence of dynamic structural effects on 1- and 2PA spectra for FMN in aqueous solution using both ADQM-Boltz. and ADQM-MD methodologies.

4.1. An All-Atom Single Structure QM Protocol Applied to Fluorescent Proteins

4.1.1. Bacteriorhodopsin

We first assess the ASQM workflow to compute the 1PA spectrum of bR (Figure S1, Supporting Information) at the sTD-DFT-xTB/GBSA(water) level of theory in comparison with experimental 1PA spectra from de Coene et al.[117] and Winder et al.[118] recorded in water. **Figure 5** presents this comparison considering

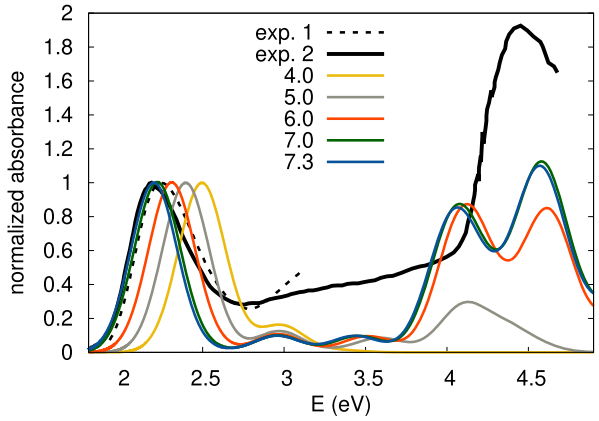

Figure 5. Experimental 1PA spectra of bR obtained by de Coene et al.[117] (exp. 1) as well as Winder et al.[118] (exp. 2) recorded in water in comparison with computed sTD-DFT-xTB/GBSA(water) spectra considering different energy thresholds (4.0, 5.0, 6.0, 7.0, and 7.3 eV). Convoluted spectra used a damping factor $\Gamma = 0.2$ eV and were shifted by -0.4 eV, allowing the sTD-DFT-xTB ($E_{thr} = 7.3$ eV) spectrum to better match experiment.

different E_{thr} values (4.0, 5.0, 6.0, 7.0, and 7.3) eV for the sTD-DFT-xTB scheme. Note that for the ASQM procedure, only the optimized structure of bR from Beaujean et al.[88] was used. **Figure 6** shows computation times and numbers of CSFs involved for sTD-DFT-xTB calculations (left panel) as a function of E_{thr} where one can observe a direct correlation between computation times and numbers of configurations. Note that the 40 min-long excited state calculation considering $E_{thr} = 7.3$ eV with the std2 program was the largest computation we were able to run without filling completely the memory of an AMD EPYC 7742 64-Core Processor (1500.0 MHz, 1024 GB RAM) computer node. Both experiments show a first maximum around 2.25 eV (552 nm)[117] and 2.19 eV (567 nm),[118] respectively. Theoretical spectra were shifted by -0.4 eV to allow the sTD-DFT-xTB spectrum computed with $E_{thr} = 7.3$ eV to best match this absorption band. The experimental spectrum from Winder et al.[118] presents two main absorption bands as also predicted by the theory. The second band energy position is actually well-reproduced by the sTD-DFT-xTB spectrum computed with the largest energy threshold when no energy shift is applied with only 0.01 eV of difference with respect to experiment. Note that at least a value of E_{thr} larger than 6 eV needs to be employed to account for this second peak providing a reasonable comparison to experiment. As expected, E_{thr} represents the spectral window. The gap between both experimental and theoretical first peak position is decreased when increasing the energy threshold going from 0.7 eV ($E_{thr} = 4$ eV) to 0.4 eV ($E_{thr} = 7.3$ eV) with respect to experiment. As both experimental spectra were normalized, we cannot discuss absolute peak intensities but only relative one. Our calculations are not able to reproduce the experimental ratio between both main absorption bands but spectral shapes remain reasonable.

Figure 7 depicts natural transition orbitals (NTOs) for the first excited state of the unoptimized retinal chromophore extracted from bR computed in vacuum with the sTD-DFT-xTB method as well as for the first and second excited states of the full protein bR at the sTD-DFT-xTB/GBSA(water) ($E_{thr} = 7.3$ eV) level of theory. The first excited state of retinal in gas phase is a pure $\pi \rightarrow \pi^*$

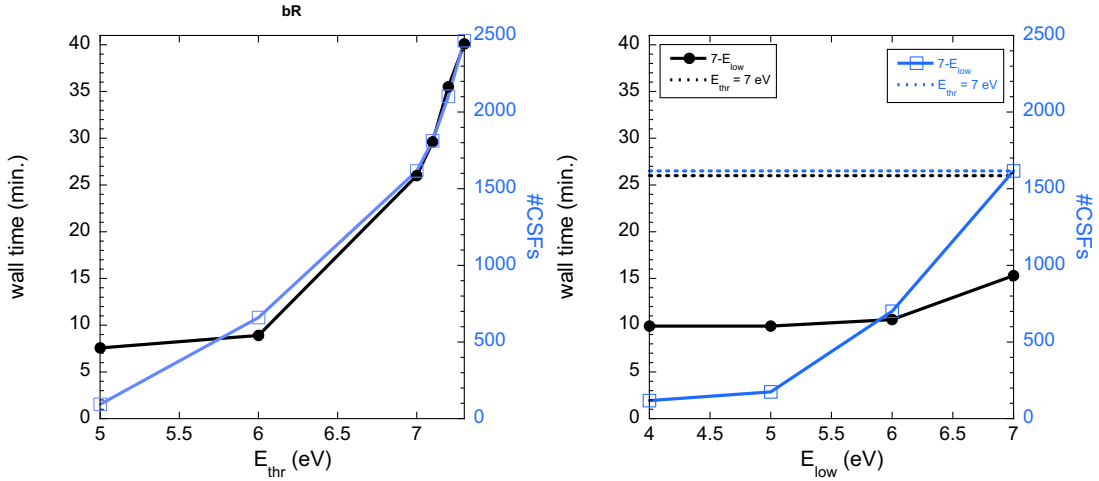

Figure 6. For bR, computation times and numbers of CSFs involved as a function of E_{thr} for sTD-DFT-xTB/GBSA(water) calculations (left panel) as well as for dt-sTD-DFT-xTB/GBSA(water) calculations (right panel) as a function of E_{low} considering $E_{high} = 7$ eV.

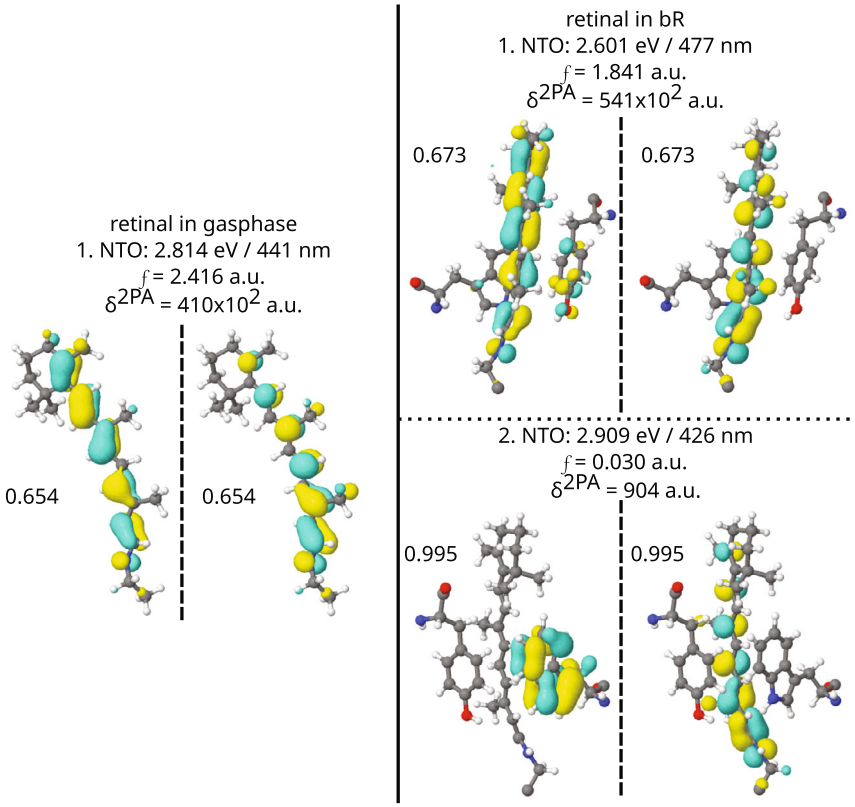

Figure 7. NTOs (left panel) computed at the sTD-DFT-xTB for the first excited state of the retinal chromophore extracted from bR as well as NTOs (right panel) computed at the sTD-DFT-xTB/GBSA(water) (both calculations with $E_{thr} = 7.3$ eV) for the first and second excited states of bR. Hole and particle NTOs are separated by dashed lines. Weights are provided for each NTO pair. The excitation energy, oscillator strength, and 2PA strength are also given for each state. Isovalue = 0.03.

excitation happening on the π-conjugated pathway of the molecule. As already mentioned in the introduction, previous studies[98–100] showed that in the protein, the retinal interacts with a nearby tyrosine (TYR185). This is confirmed by our NTO analysis. For the first excited state, we have the same $\pi \rightarrow \pi^*$ excitation as for the retinal in gas phase but with contributions from the tyrosine in the hole NTO. This small charge transfer (CT) contribution stabilizes this electronic transition energy by 0.21 eV and the oscillator strength is reduced from 2.416 to 1.841. Interestingly, the NTOs analysis (not shown here) shows that both higher energy peaks are composed of collection of $\pi \rightarrow \pi^*$ transitions located on tryptophans (TRP80, 137, 138, and 198) from the protein. Obviously, such transitions can only be captured by accounting for the whole protein into the excited state calculation. This result advocates for the use of the ASQM methodology to compute 1PA spectra of proteins.

Figure 8 compares the sTD-DFT-xTB/GBSA(water) spectrum computed with $E_{thr} = 7.0$ eV to dt-sTD-DFT-xTB/GBSA(water) ones computed with $E_{high} = 7.0$ eV and E_{low} values of 4.0, 5.0, and 6.0 eV. To reproduce the second larger absorption band around 4 eV from the sTD-DFT-xTB(7.0 eV) calculation, we need to use at least an energy threshold of 6.0 eV for the protein surroundings because it involves $\pi \rightarrow \pi^*$ transitions located on tryptophans. Note that dt-sTD-DFT spectra were shifted by a larger energy shift of −0.6 eV to match the lowest-energy peak. Thus, the 0.2 eV difference of energy observed in Figure 8 for the second band between

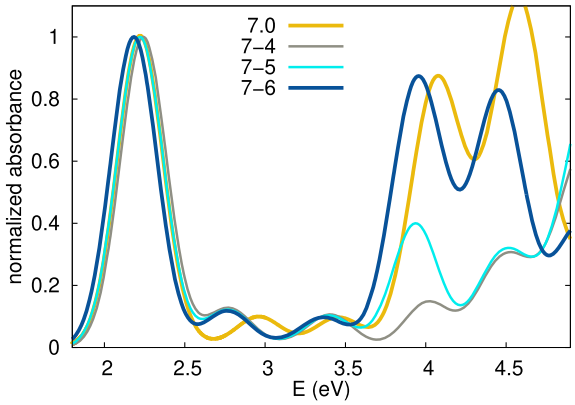

Figure 8. For bR, sTD-DFT-xTB/GBSA(water) 1PA spectrum computed with $E_{thr} = 7.0$ eV in comparison with dt-sTD-DFT-xTB/GBSA(water) ones computed with $E_{high} = 7.0$ eV and E_{low} values of 4.0, 5.0, and 6.0 eV. Convoluted spectra used a damping factor $\Gamma = 0.2$ eV and were shifted by −0.4 eV for sTD-DFT-xTB and −0.6 eV for dt-sTD-DFT-xTB.

both sTD-DFT-xTB (7.0 eV) and 7-6 spectra is artificial as both excitation energies are 4.48 and 4.55 eV, respectively. Figure 6 shows computation times and numbers of CSFs involved for sTD-DFT-xTB (left panel) and dt-sTD-DFT-xTB calculations (right panel) as a function of E_{low} (see Table S2, Supporting Information for more details). Using 1351 CSFs, the 7-6 dt-sTD-DFT-xTB calculation took only 11 min, while the sTD-DFT-xTB(7.0 eV) one, using 1616 CSFs, took

26 min, reducing by more than a factor of two the computation time without affecting much the global shape of the spectrum. Note that the 7–4 spectrum is already reproducing well the sTD-DFT-xTB(7.0 eV) one for (shifted) excitation energies below 3.5 eV but the extra cost to perform 7–6 dt-sTD-DFT-xTB calculation is negligible with respect to the 7–4 one. While leading to the same configuration space of 1616 CSFs, the 7–7 calculation took less time (15 min) than the sTD-DFT-xTB(7.0 eV) one (26 min) as less CSFs were probed by the perturbative approach to determine S-CSFs due to the restricted active space defined by Equation (13).

Second, we assess the ASQM approach to evaluate the 2PA spectrum of bR with the sTD-DFT-xTB/GBSA(water) with respect to the experiment. **Figure 9** presents the experimental 2PA spectrum for bR recorded in water by Birge et al.[102] in 1990 in comparison with computed sTD-DFT-xTB/GBSA(water) spectra considering different energy thresholds (5.0, 6.0, 7.0, 7.1, 7.2, and 7.3 eV). Note that according to data from Birge et al.[102] (see Table 1 in ref. [102]), the experimental spectrum was not corrected for background voltage, and all 2PA cross-sections were contaminated by 1PA signals. To improve the comparison with respect to this rather old experimental 2PA spectrum, we corrected arbitrarily the baseline by shifting the minimum error bar of the lowest-energy point at 0 GM for all experimental 2PA cross sections. Note too that the 2PA spectrum was recorded for a very narrow 1 eV-wide spectral range. This 2PA spectrum shows a distinct peak at 2.27 eV with a maximum of 191 (±38) GM, a shoulder around 2.5 eV, and the onset of a another peak around 3 eV. Computed 2PA spectra were shifted by −0.4 eV as 1PA ones. Our calculations reproduce well the global shape of the experimental 2PA spectrum even if the shoulder of the first excited state is not reproduced. It could be due to vibronic effects that are not accounted for by our calculations or to the second excited state for which the 2PA cross-section could be underestimated. Figure 7 depicts the NTO pair for the second excited state of bR computed at the sTD-DFT-xTB/GBSA(water) (E_{thr} = 7.3 eV) level of theory, showing a CT transition at 2.91 eV from a tryptophan (TRP86) to the retinal chromophore with δ^{2PA} = 904 a.u. Note that for the first excited state, the 2PA cross section is enhanced from the gas phase retinal calculation from 410 to 541 × 10^2 GM for the full protein thanks to the nearby tyrosine giving the CT character to the transition. The onset of the third band is also computed. Because of the baseline problem of the experiment and that 2PA bands were not filtered from 1PA, quantitative comparisons on 2PA cross sections with respect to experiment are especially challenging. However, computed and experimental cross sections are more or less of the same order of magnitude.

Regarding the impact of the energy threshold on calculations, as we observed for 1PA spectra, the energy gap between the first experimental 2PA band and the sTD-DFT-xTB computed one decreases when increasing this threshold, while the rest of the spectrum is not changing much. **Figure 10** compares the experimental spectrum recorded by Birge et al.[102] in water with the computed sTD-DFT-xTB/GBSA(water) spectrum considering an energy threshold of 7.0 eV as well as dt-sTD-DFT-xTB/GBSA(water) spectra with 6–4, 6–5, 7–4, 7–5, and 7–6 values of E_{high}–E_{low} (see Figure S2, Supporting Information, for comparison of sTD-DFT and dt-sTD-DFT for more energy thresholds). As for 1PA, a larger energy red-shift of −0.6 eV was applied for dt-sTD-DFT-xTB 2PA spectra. The influence of chosen energy threshold on dt-sTD-DFT-xTB 2PA calculations remains negligible. All dt-sTD-DFT-xTB 2PA spectra show a first distinct peak around 2.2–2.3 eV fitting well to experiment. The main difference between sTD-DFT and dt-sTD-DFT 2PA spectra is a decreased energy difference between first and second main peaks for dt-sTD-DFT results with respect to sTD-DFT ones. The first excitation at around 2.2. eV has a somewhat smaller magnitude with the dt-sTD-DFT scheme with respect to the sTD-DFT one (7–6 eV: 120 GM vs. 7.0 eV: 142 GM). Nevertheless, the experimental spectrum is rather well reproduced by both methods, even considering small energy thresholds, for example, 6.0 eV or 6–4 eV.

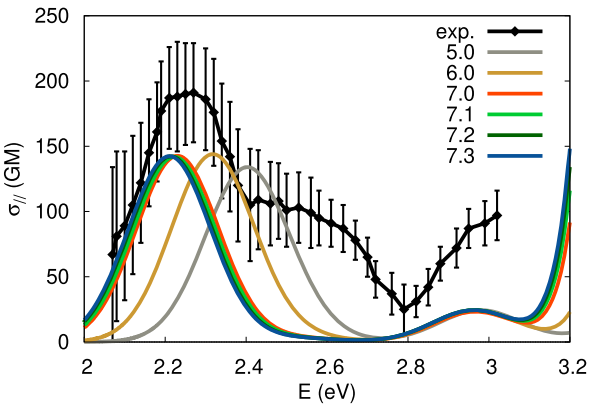

Figure 9. Experimental 2PA spectrum of bR recorded by Birge et al.[102] in water in comparison with computed sTD-DFT-xTB/GBSA(water) spectra considering different energy thresholds (5.0, 6.0, 7.0, 7.1, 7.2, and 7.3 eV). Convoluted spectra used a damping factor Γ = 0.1 eV and were shifted by −0.4 eV, allowing the sTD-DFT-xTB (E_{thr} = 7.3 eV) spectrum to better match experiment.

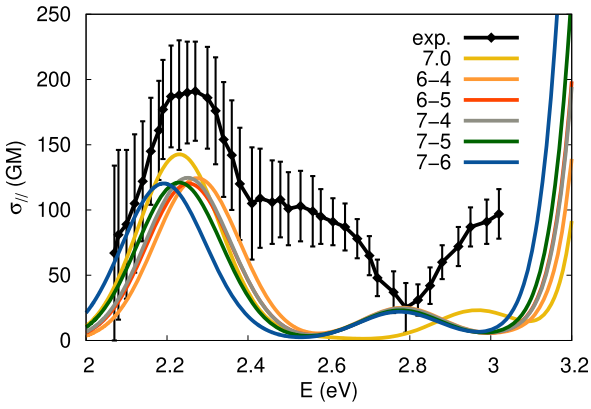

Figure 10. Experimental 2PA spectrum of bR recorded by Birge et al.[102] in water in comparison with the computed sTD-DFT-xTB/GBSA(water) spectrum considering an energy threshold 7.0 eV as well as dt-sTD-DFT-xTB/GBSA(water) spectra with 6–4, 6–5, 7–4, 7–5, and 7–6 values of $E_{high} - E_{low}$. Convoluted spectra used a damping factor Γ = 0.1 eV and were shifted by −0.4 eV for sTD-DFT-xTB and −0.6 eV for dt-sTD-DFT-xTB.

4.1.2. iLOV

For iLOV[105] (Figure S3, Supporting Information), we further assess the performance of the ASQM workflow to evaluate 1- and 2PA of such large systems. Again, only the optimized structure of iLOV from Beaujean et al.[88] was used. **Figure 11** presents 1PA spectra recorded in water by Ran et al.[108] as well as Torra et al.[109] in comparison with computed sTD-DFT-xTB/GBSA(water) spectra considering different energy thresholds (6.0, 7.0, 8.0, 9.0, and 9.6 eV). Note that no energy shifts were applied to calculated spectra and that they were normalized to match the lowest-energy band at 2.79 eV (445 nm). Both experimental 1PA spectra show three main absorption bands. The first peak presents a clearly resolved vibronic structure, similar to the one found by List et al.[4] emerging because of the protein environment as for miniSOG (see Figure 2). The left panel of **Figure 12** shows on a logarithmic scale the computational time and number of selected CSFs as a function of E_{thr} for the sTD-DFT-xTB calculations. On an AMD EPYC 7742 64-Core Processor (1500.0 MHz, 1024 GB RAM) computer node, the largest sTD-DFT-xTB excited state calculation we were able to run was with $E_{thr} = 9.6$ eV considering 46 027 CSFs and took a bit less than 41 h. Using an energy threshold of 6.0 eV reduces the computation time to only 2 min. Comparing the sTD-DFT-xTB/GBSA(water) ($E_{thr} = 9.6$ eV) spectrum to experiments shows that the three main peaks are reproduced. For the first peak, the theoretical excitation energy is only 0.08 eV away from the experimental maximum and for the third peak, we have 0.07 eV of difference. Note that if only the chromophore of iLOV was accounted for, the first excited state energy would be red-shifted by 0.68 eV with respect to the calculation accounting for the whole protein. The second peak in Figure 11 is blue-shifted with respect to experiment by 0.46 eV. This blue shift is related to the underlying isoalloxazine structure and its planarity. During the ONIOM optimization,[88] the original planar ring system of the isoalloxazine from iLOV crystal structure was slightly bended. As already mentioned in the introduction, the 1PA spectrum of FMN is very sensitive to the planarity of the isoalloxazine ring.[76,110,111,113] To test this further, we computed the 1PA spectrum of iLOV while keeping the isoalloxazine flat. Position and intensities of the second and third excitations show significant changes, while the first signal remains unaltered (data not shown here). When increasing E_{thr}, computed 1PA spectra are red-shifted and intensities of the second and third bands are increased, better matching experiments.

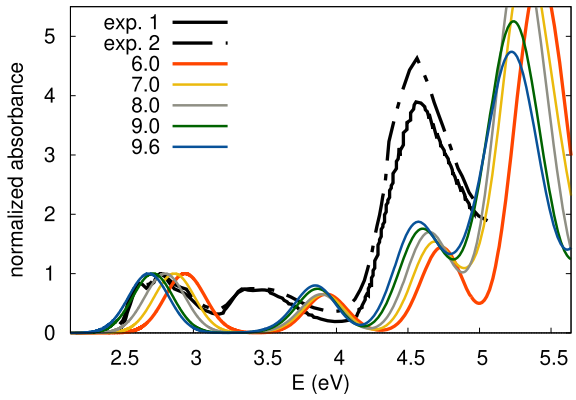

Figure 11. Experimental 1PA spectra of iLOV obtained by Ran et al.[108] (exp. 1) as well as Torra et al.[109] (exp. 2) recorded in water in comparison with computed sTD-DFT-xTB/GBSA(water) spectra considering different energy thresholds (6.0, 7.0, 8.0, 9.0, and 9.6 eV). Convoluted spectra used a damping factor $\Gamma = 0.2$ eV and no energy shifts were applied.

Figure 13 presents NTOs for the three main excitations of the 1PA spectrum computed at the sTD-DFT-xTB/GBSA(water) ($E_{thr} = 8.0$ eV) level of theory. The first excited state at 2.80 eV presents a $\pi \rightarrow \pi^*$ transition located on the isoalloxazine ring. No CT character is observed from nearby residues explaining why the signal is almost not altered by the chromophore direct environment (see Figure 2). For the second excited state at 3.65 eV, the hole NTO is a linear combination of *n* orbitals located on hetero-atoms as well as σ orbitals from the isoalloxazine ring with a small contribution from a *P* orbital on the oxygen from a nearby asparagine. The particle NTO is the same π^* orbital as for the first excited state. This transition has a small oscillator

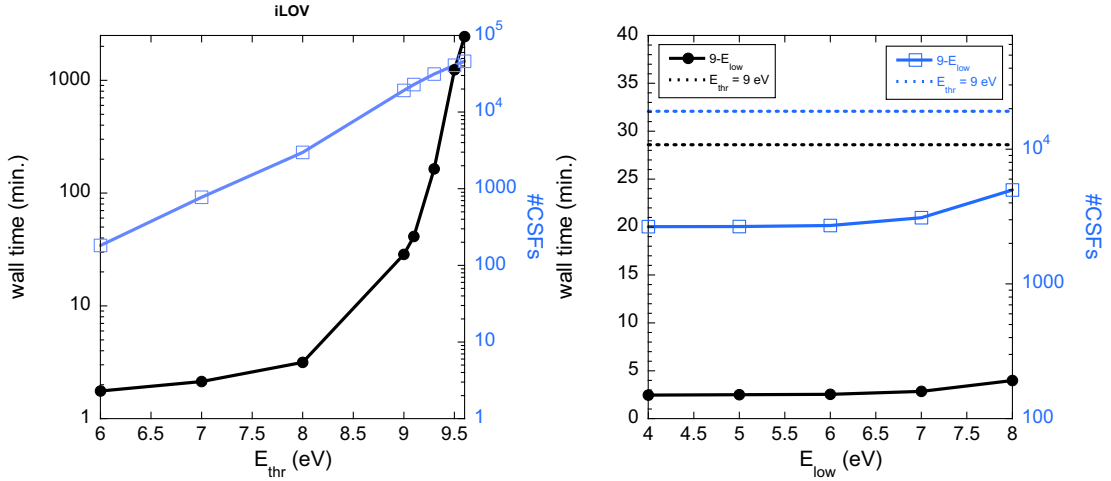

Figure 12. For iLOV, computation times and numbers of CSFs involved as a function of E_{thr} for sTD-DFT-xTB calculations (left panel) as well as for dt-sTD-DFT-xTB calculations (right panel) as a function of E_{low} considering $E_{high} = 9$ eV.

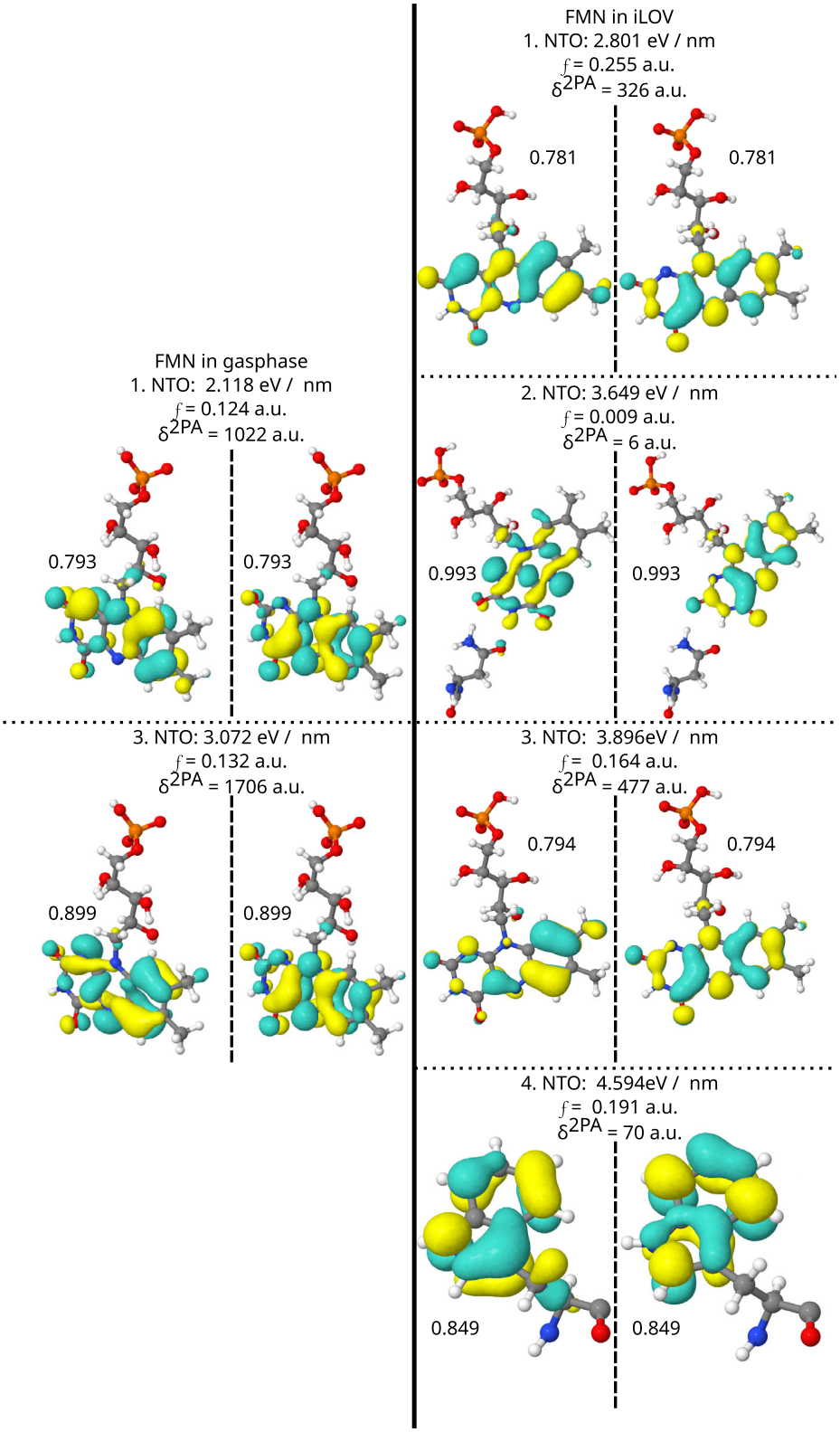

Figure 13. NTOs (left panel) computed at sTD-DFT-xTB(vacuum) for the first and third excited states of FMN chromophore extracted from iLOV as well as NTOs (right panel) computed at the sTD-DFT-xTB/GBSA(water) for the first, second, third, and fourth excited states of iLOV (both calculations with $E_{thr} = 8.0$ eV). Hole and particle NTOs are separated by dashed lines. Weights are provided for each NTO pair. The excitation energy, oscillator strength, and 2PA strength are also given for each state. Isovalue = 0.03.

strength of 0.01 and does not contribute much to the second absorption band. The third excited state is the main contribution to this second peak with an oscillator strength of 0.16. It is another $\pi \rightarrow \pi^*$ transition located on the isoalloxazine ring. More interestingly, the forth excited state is the result of a $\pi \rightarrow \pi^*$ excitation located on a tryptophan of the protein and

thus not involving the chromophore. As for bR and also observed for PYP by one of us,[93] an AQM protocol is needed to compute properly absorption spectra of fluorescent proteins, especially to account for $\pi \rightarrow \pi^*$ transitions located on tryptophan residues that impact the high energy part of spectra.

Figure 14 compares experimental 1PA spectra to computed ones with both sTD-DFT-xTB/GBSA(water) ($E_{thr} = 9.0$ eV) and dt-sTD-DFT-xTB/GBSA(water) (9–4, 9–5, 9–6, 9–7, and 9–8 values of $E_{high} - E_{low}$) schemes (see Figure S4, Supporting Information, for a more detailed comparison of sTD-DFT and dt-sTD-DFT considering more energy thresholds). The right panel of Figure 12 presents computation times and numbers of CSFs as a function of E_{low} for dt-sTD-DFT-xTB calculations considering $E_{high} = 9.0$ eV (see Table S3, Supporting Information for more details). While 9–4 and 9–5 dt-sTD-DFT-xTB computed spectra failed to reproduce correctly the third absorption band with respect to the sTD-DFT-xTB spectrum, using a E_{low} value of 6.0 eV seems good enough to replicate the sTD-DFT-xTB/GBSA(water) ($E_{thr} = 9.0$ eV) spectrum. As already observed for bR, a value of $E_{low} \geq 6$ eV is necessary to properly cope with tryptophan $\pi \rightarrow \pi^*$ transitions. The configuration space is reduced from 19 165 with $E_{thr} = 9.0$ eV to 2716 CSFs with 9-6 values of $E_{high} - E_{low}$, reducing the computation time by a factor of 10. Note that for sTD-DFT-xTB calculations, it was not possible to use larger E_{thr} value than 9.6 eV on an AMD EPYC 7742 64-Core Processor (1500.0 MHz, 1024 GB RAM) computer node, but we were able to run dt-sTD-DFT-xTB calculations with $E_{high} = 10.0$ eV, increasing further the space of CSFs for the chromophore (see Figure S4, Supporting Information, lower panel). Using at least $E_{low} = 6.0$ eV, dt-sTD-DFT-xTB spectra are very similar to the sTD-DFT-xTB ($E_{thr} = 9.6$ eV) one.

Following the discussions on the 1PA of iLOV, we further assess the ASQM approach to evaluate the 2PA spectrum of iLOV with the sTD-DFT-xTB/GBSA(water). **Figure 15** presents the experimental 2PA spectrum of iLOV recorded in water by Homans et al.[119] in comparison with computed sTD-DFT-xTB/GBSA(water) spectra considering different energy thresholds (6.0, 7.0, 8.0, 9.0, 9.3, 9.5, and 9.6 eV). The experimental 2PA spectrum of iLOV[119] is recorded for a small spectral range from 2.55 to 3.35 eV and shows a maximum at around 2.9 eV (≈0.9 GM). The onset of a second peak around 3.35 eV (≈1.5 GM) was also recorded. Computed sTD-DFT-xTB spectra are slightly shifted (±0.1 eV) with respect to experiment but the comparison is striking, very well reproducing the 2PA cross-section of first 2PA band (with $E_{thr} = 9.6$ eV, $\sigma^{2PA} = 0.70$ GM). All sTD-DFT-xTB spectra reproduce the first experimental peak in position and intensity, while the second peak (which is not clearly resolved in the experiment) is always blue-shifted, as we also observed for computed 1PA spectra (vide supra). The variance between spectra computed with different thresholds is small; therefore, it is possible to go as low as $E_{thr} = 6.0$ eV for this system because of the small energy range considered by the experiment. For the dt-sTD-DFT-xTB spectra (see **Figure 16** and S4, Supporting Information), a similar picture emerges. Even the smallest tested 9 − 4 ($E_{high} - E_{low}$) thresholds lead to an excellent comparison with respect to experiment and the calculation took less than 3 min.

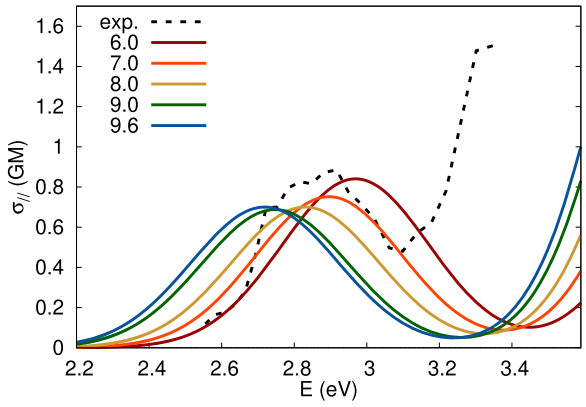

Figure 15. Experimental 2PA spectrum of iLOV recorded by Homans et al.[119] in water in comparison with computed sTD-DFT-xTB/GBSA(water) spectra considering different energy thresholds (6.0, 7.0, 8.0, 9.0, and 9.6 eV). Convoluted spectra used a damping factor Γ = 0.2 eV and no energy shifts were applied.

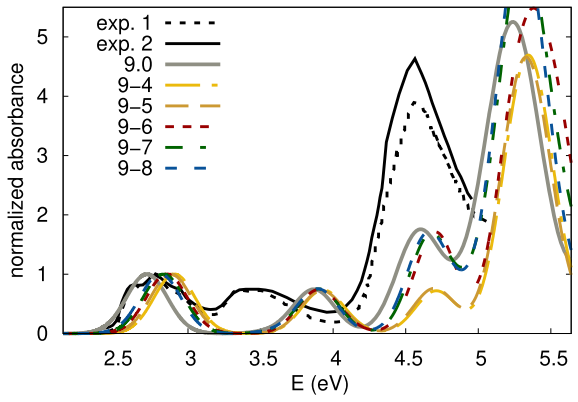

Figure 14. Experimental 1PA spectra of iLOV obtained by Ran et al.[108] (exp. 1) as well as Torra et al.[109] (exp. 2) recorded in water in comparison with computed sTD-DFT-xTB/GBSA(water) ($E_{thr} = 9.0$ eV) spectrum as well as dt-sTD-DFT-xTB/GBSA(water) ones with 9–4, 9–5, 9–6, 9–7, and 9–8 values of $E_{high} - E_{low}$. Convoluted spectra used a damping factor Γ = 0.2 eV and no energy shifts were applied.

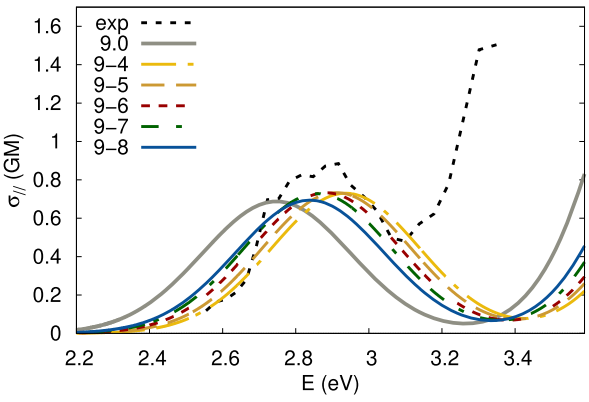

Figure 16. Experimental 2PA spectrum of iLOV recorded by Homans et al.[119] in water in comparison with the computed sTD-DFT-xTB/GBSA(water) spectrum considering an energy threshold 9.0 eV as well as dt-sTD-DFT-xTB/GBSA(water) spectra with 9–4, 9–5, 9–6, 9–7, and 9–8 values of $E_{high} - E_{low}$. Convoluted spectra used a damping factor Γ = 0.2 eV and no energy shifts were applied.

As intermediate conclusions, the ASQM methodology proposed here provided excellent comparisons for the 1- and 2PA spectra of bR and iLOV with respect to experiment. For both proteins, an ASQM workflow was necessary to describe absorption bands involving tryptophan residues, confirming what was already shown by one of us[93] for PYP. Note that we are at edge of what is computationally feasible with quantum chemistry for such systems. However, including dynamic structural effects could improve further comparisons with respect to experiment as we will show in the next section for FMN in aqueous solution.

4.2. An All-Atom Dynamic Structure QM Protocol Applied to the Flavin Mononucleotide in Water Solution

We could still argue that for FPs, because of the network of non-covalent interactions (including H-bonds), the environment of the chromophore remains rigid enough to justify the use of an ASQM procedure. However, subtle distortions or new configurations of the chromophore can still change the 1- and 2PA spectra dramatically. For solvated dyes, an ASQM approach is certainly not sufficient. Dynamic structural effects need to be included in. In the next section, both ADQM-Boltz. and ADQM-MD approaches are assessed to compute the 1- and 2PA of FMN in aqueous solution. Before applying these ADQM schemes, 1- and 2PA spectra computed in vacuum at the sTD-DFT-xTB level are discussed and compared to reference RI-CC2 ones as well as using implicit solvent models.

Figure 2 presents experimental 1PA spectra of FMN recorded in water, methanol, iLOV, and miniSOG. Note that all spectra were normalized on the first absorption peak. As so, we cannot assess the impact of the surroundings on the molar absorption coefficient. Independent of the environment, the lowest-energy peak always arises around 2.79 eV (445 nm). In the protein surroundings, this absorption band is more vibrationally resolved. Going from water (3.31 eV or 375 nm) to methanol or proteins (3.49 eV or 355 nm) environments, the second signal is shifted by 0.18 eV (20 nm). The energy position of this second absorption band seems dependent on the polarity of the environment. Both proteins influence this peak similarly as methanol, although it was suggested that the dielectric constant of such protein environment (9.2) is way lower than the one from methanol (32.6).[78,178] The intensity of the third peak in the UV-C region (4.59 eV or 270 nm), traced back to be influenced by the environment,[108] decreases in water compared to FMN in protein. As we discussed in the previous section, this band is clearly impacted by the absorption of tryptophan residues from the protein, explaining this trend.

Figure 17 compares 1PA spectra of FMN computed in gas phase at RI-CC2/aug-cc-pVDZ and sTD-DFT-xTB levels of theory to experimental spectra recorded in water. It also reports the average sTD-DFT-xTB spectrum on uncorrelated snapshots from a GFN2-xTB MD that was run in vacuum for 11 ns. Note that the 1PA spectrum from Antill et al.[106] was normalized; therefore, we extrapolated it to match the extinction coefficient of the first absorption band of the 1PA spectrum recorded in water by List et al.[4] The first experimental absorption band around

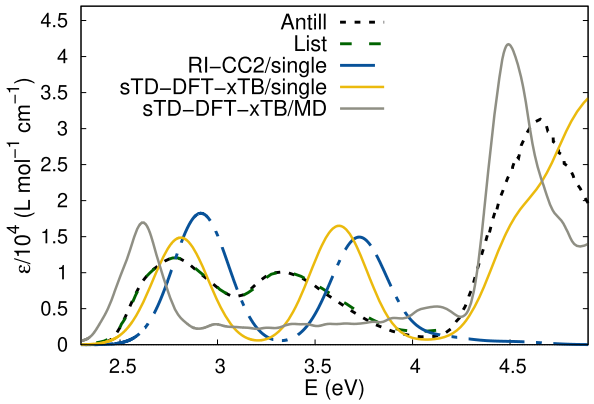

Figure 17. Experimental 1PA spectra of FMN recorded in water by Antill et al.[106] as well as List et al.[4] in comparison with computed RI-CC2/aug-cc-pVDZ//r^{2}SCAN-3c and sTD-DFT-xTB//r^{2}SCAN-3c (E_{thr} = 16.0 eV) spectra in vacuum as well as the average sTD-DFT-xTB spectrum on uncorrelated snapshots of a GFN2-xTB 11 ns-long MD in gas phase. Convoluted spectra used a damping factor Γ = 0.2 eV, while a damping factor Γ = 0.04 eV was used for spectra computed for each MD snapshot. No energy shifts were applied.

2.79 eV is rather well-reproduced by both RI-CC2 and sTD-DFT-xTB calculations. With respect to RI-CC2, the sTD-DFT-xTB excitation energy is red-shifted by 0.1 eV, and the oscillator strength is 19% smaller. Considering the second absorption band for which experiment showed its sensitivity to the surroundings,[78,178] as expected, large blue shifts of 0.39 and 0.28 eV with respect to experiment are observed at both RI-CC2 and sTD-DFT-xTB levels of theory, respectively. The average sTD-DFT-xTB spectrum computed from uncorrelated snapshots of a GFN2-xTB MD shows a complete compression of the second absorption band. Globally, RI-CC2 and sTD-DFT-xTB 1PA spectra in vacuum are very similar, justifying the usage of the sTD-DFT-xTB method to compute 1PA spectra for FMN.

Let us assess the ADQM-Boltz. approach to simulate the 1PA spectrum of FMN in aqueous solution. Dynamic structural effects were accounted for implicitly by Boltzmann-averaging 1PA spectra computed at the sTD-DFT-xTB/GBSA(water) for the 23 lowest-energy conformers obtained at the r^{2}SCAN-3c level of theory following the CENSO[124,125] protocol. For the sake of completeness, comparisons were made using the r^{2}SCAN-3c lowest-energy conformer to compute excited states with both RI-CC2 and sTD-DFT-xTB levels using implicit solvent models (DCOSMO-RS and GBSA, respectively). **Figure 18** compares experimental 1PA spectra recorded in water[4,106] to spectra computed at these levels of theory as well as the Boltzmann-averaged sTD-DFT-xTB/GBSA(water) spectrum obtained by following the ADQM-Boltz. workflow. It becomes very clear that either using the lowest-energy conformer only or the ensemble of conformers barely show any difference on the sTD-DFT-xTB 1PA spectra. Note that the lowest-energy conformer has only a Boltzmann weight of around 25%. The position of the first energy band at 2.79 eV is well-reproduced by both RI-CC2 and sTD-DFT-xTB methods with only 0.08 and 0.04 eV of difference, respectively. Both RI-CC2 and sTD-DFT-xTB schemes overestimate the intensity by a factor of 2.1 and 1.4. The second peak is blue-shifted by both methods

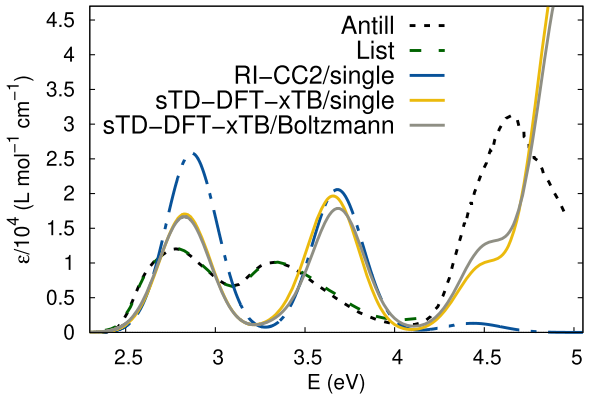

Figure 18. Experimental 1PA spectra of FMN recorded in water by Antill et al.[106] as well as List et al.[4] in comparison with computed RI-CC2/aug-cc-pVDZ//r^{2}SCAN-3c and sTD-DFT-xTB//r^{2}SCAN-3c ($E_{thr} = 16.0$ eV) spectra using implicit solvent models (DCOSMO-RS and GBSA, respectively) for water as well as the Boltzmann-averaged sTD-DFT-xTB/GBSA(water) spectrum for the conformer ensemble obtained at room temperature with the r^{2}SCAN-3c method. Convoluted spectra used a damping factor $\Gamma = 0.2$ eV. No energy shifts were applied.

(RI-CC2/single: 3.68 eV and sTD-DFT-xTB: 3.66 eV) compared to the experimental peak maximum at 3.31 eV. The intensity of the second peak is overestimated by both methods with respect to experiment. Nevertheless, RI-CC2 reproduces the trend between the first and second peak intensities, while sTD-DFT-xTB inverts this relation. These results are not encouraging to employ the ADQM-Boltz. workflow for this very case.

The ADQM-MD protocol is applied to compute the 1PA spectrum of FMN, explicitly accounting for water molecules and dynamic structural effects. **Figure 19** compares experimental 1PA spectra recorded in water[4,106] to average dt-sTD-DFT/GBSA(water) spectra on uncorrelated snapshots extracted from 1 ns-long GFN2-xTB MD simulations accounting explicitly for solvent molecules using solvation spheres of increasingly large radii of 7, 9, 11 and 12 Å (see Figure S6, Supporting Information) as well

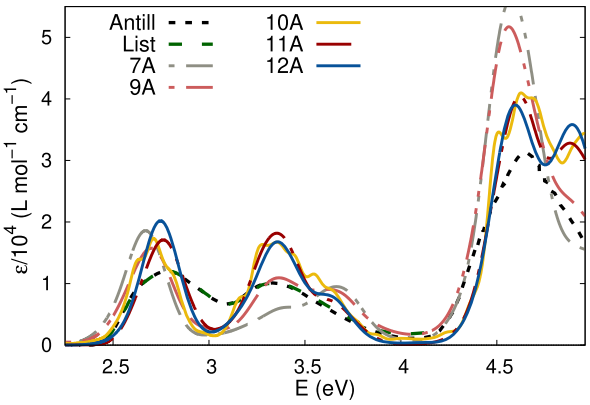

Figure 19. Experimental 1PA spectra of FMN recorded in water by Antill et al.[106] as well as List et al.[4] in comparison with computed average dt-sTD-DFT-xTB/GBSA(water) spectra on uncorrelated snapshots extracted from 1 ns-long GFN2-xTB MD simulations accounting explicitly for solvent molecules using solvation spheres of increasingly large radii of 7, 9, 11, and 12 Å. Note that when considering a sphere of 10 Å, a longer 3 ns-long GFN2-xTB MD simulation was run. Convoluted spectra used a damping factor $\Gamma = 0.1$ eV. No energy shifts were applied.

as from a 3 ns-long GFN2-xTB MD simulation considering a solvation sphere of 10 Å. In consequence, different amounts of water molecules (124,235,307,386 and 483 for spheres of 7, 9, 10, 11, and 12 Å, respectively) were considered to assess the convergence of the ADQM-MD procedure. Note that spectra were computed with the dt-sTD-DFT/GBSA(water) method to reduce the computational cost using $E_{high} - E_{low}$ values of 11–8 eV, except for the dt-sTD-DFT-xTB(12 Å)/GBSA(water) for which we used 10–8 eV. When considering a sphere of 12 Å, the sTD-DFT-xTB calculation with $E_{thr} = 10$ eV took 46 min (on the same computer node specified above), while the dt-sTD-DFT-xTB with $E_{high} - E_{low}$ values of 10–8 eV took only 16 min without modifying much the computed spectrum, justifying this choice. The comparison with respect to experiment is more than striking, knowing that no energy shifts were applied to MD-average dt-sTD-DFT-xTB spectra. While simulated spectra are still arbitrary damped, energy position, intensities, and broadening of the three main experimental peaks are well-reproduced. Considering the dt-sTD-DFT-xTB(12 Å)/GBSA(water) MD-average spectrum, for example, the first absorption band occurs at 2.75 eV (451 nm, exp: 2.79 eV or 445 nm), and the second peak at 3.36 eV (369 nm, exp: 3.31 eV or 375 nm) with intensities of 20 194 and 16 780 L mol^{-1} cm^{-1} (exp: 12 043 and 10 122 L mol^{-1} cm^{-1}), respectively. The reproduction of the second experimental absorption band is very sensitive to the size of the solvation sphere included into the ADQM-MD procedure. At least a solvation sphere of 9 Å is needed to describe properly this second peak, but for better describing the third absorption band, a minimum 10 Å solvation sphere should be used. Only the ADQM-MD was able to account for the very specific broadening of the second absorption band. All in all, the ADQM-MD procedure that accounts explicitly for the environment and explicitly for dynamic structural effects correctly described the first three main absorption bands of the 1PA spectrum of FMN in aqueous solution.

To understand what is responsible for the broadening of the second absorption band, NTOs were calculated from two representative snapshots of the GFN2-xTB MD. **Figure 20** presents NTOs computed at dt-sTD-DFT-xTB/GBSA(water) (10 − 8 eV) for two excited states of FMN in explicit water (12 Å) that contribute to the second absorption band. We already saw for iLOV that the second absorption band is due to a $\pi \rightarrow \pi^*$ transition on the isoalloxazine ring. For FMN in water, the dynamic nature of the environment directly impacts this transition. For both NTO transitions, the hole orbital is a linear combination of the π orbital from the isoalloxazine ring and nearby *n* orbitals from oxygen located on water molecules as well as from the tail of FMN chromophore. Particle orbitals from both NTOs are mostly π^* orbitals located on the isoalloxazine ring. In both cases, the $\pi \rightarrow \pi^*$ transition presents a CT character modulated by the interaction of the chromophore with its nearby surroundings. Thus, the dynamic interaction of the HPO_4^- group as well as water molecules with the isoalloxazine ring is crucial to correctly reproduce the second absorption band and its broadening. This advocates for the use of the ADQM-MD scheme.

Let us now focus on 2PA of FMN in water, assessing both ADQM schemes for computing such spectra. As before, we firstly

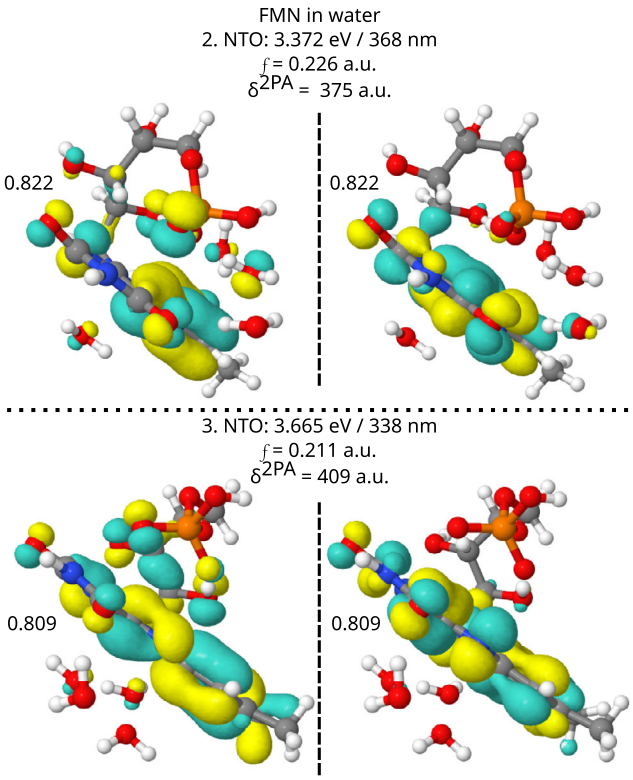

Figure 20. From two different snapshots of the GFN2-xTB MD, NTOs computed at dt-sTD-DFT-xTB/GBSA(water) (10 − 8 eV) for two excited states of FMN in explicit water (12 Å) that contribute to the second absorption band. Hole and particle NTOs are separated by dashed lines. Weights are provided for each NTO pair. The excitation energy, oscillator strength, and 2 PA strength are also given for each state. Isovalue = 0.03.

discuss FMN results obtained in vacuum. **Figure 21** compares two experimental 2PA spectra of FMN recorded in water[4,119] to computed RI-CC2/aug-cc-pVDZ and sTD-DFT-xTB (E_{thr} = 16.0 eV) 2PA spectra calculated in vacuum as well as the average sTD-DFT-xTB spectrum on uncorrelated snapshots from a GFN2-xTB 11 ns-long MD in gas phase. Both experimental 2PA spectra were recorded for a narrow energy range. Combining both experimental data leads to an energy window larger by 0.5 eV than for iLOV. The experimental 2PA spectra (Figure 21) feature two main peaks around 2.88 and 3.43 eV. Comparing them to the RI-CC2 spectrum, we observe a good agreement for the first experimental peak energy position with only 0.04 eV of difference with respect to experiment but 8% of overestimation for the 2PA cross section. The 2PA cross section of the second absorption band is nine times larger than the experimental one (1.77 GM) while slightly underestimated at the sTD-DFT-xTB level. The average sTD-DFT-xTB spectrum computed from uncorrelated snapshots of a GFN2-xTB MD improves the comparison with respect to experiment for the second 2PA band but the onset of the band continues to climb at higher energies. Note that the sTD-DFT-xTB calculation on the optimized r^2SCAN-3c geometry took only 4 s, while the RI-CC2 calculation needed 130 h (wall time) for the computation of both ground and excited states as well as 2PA strengths.

As 1PA and 2PA spectra are involving same transitions, we can expect that the ADQM-Boltz. shall not be particularly helpful to cope with dynamic structural effects. **Figure 22** confirms this, comparing experimental 2PA spectra recorded in water[4,119] to spectra computed at these levels of theory as well as the Boltzmann-averaged sTD-DFT-xTB/GBSA(water) spectrum obtained by following the ADQM-Boltz. workflow.

Finally, the ADQM-MD workflow is applied to evaluate the 2PA of FMN in water. The same procedure is followed as for 1PA. **Figure 23** compares experimental 2PA spectra of FMN recorded in water by Homans et al.[119] as well as List et al.[4] to average dt-sTD-DFT-xTB/GBSA(water) spectra on uncorrelated snapshots extracted from 1 ns-long GFN2-xTB MD simulations accounting explicitly for solvent molecules using solvation spheres of increasingly large radii of 7, 9, 11, and 12 Å as well as from a 3 ns-long GFN2-xTB MD simulation considering a solvation sphere of 10 Å. For 2PA, the comparison with respect to

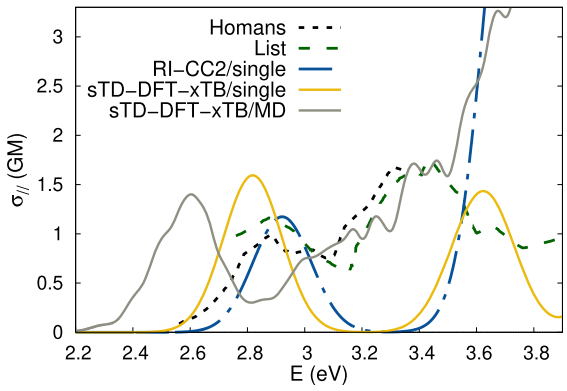

Figure 21. Experimental 2PA spectra of FMN recorded in water by Homans et al.[119] as well as List et al.[4] in comparison with computed RI-CC2/aug-cc-pVDZ//r^2SCAN-3c and sTD-DFT-xTB//r^2SCAN-3c (E_{thr} = 16.0 eV) spectra in vacuum as well as the average sTD-DFT-xTB spectrum on uncorrelated snapshots from a GFN2-xTB 11 ns-long MD in gas phase. Convoluted spectra used a damping factor Γ = 0.1 eV, while a damping factor Γ = 0.02 eV was used for spectra computed for each MD snapshot. No energy shifts were applied.

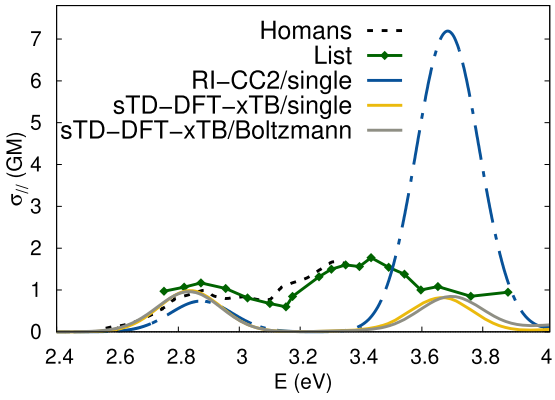

Figure 22. Experimental 2PA spectra of FMN recorded in water by Homans et al.[119] as well as List et al.[4] in comparison with computed RI-CC2/aug-cc-pVDZ//r^2SCAN-3c and sTD-DFT-xTB//r^2SCAN-3c (E_{thr} = 16.0 eV) spectra using implicit solvent models (DCOSMO-RS and GBSA, respectively) for water as well as the Boltzmann-averaged sTD-DFT-xTB/GBSA(water) spectrum for the conformer ensemble obtained at room temperature with the r^2SCAN-3c method. Convoluted spectra used a damping factor Γ = 0.1 eV. No energy shifts were applied.

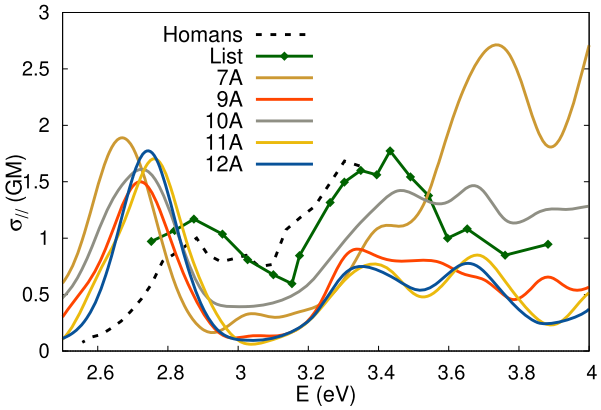

Figure 23. Experimental 2PA spectra of FMN recorded in water by Homans et al.[119] as well as List et al.[4] in comparison with computed average dt-sTD-DFT-xTB/GBSA(water) spectra on uncorrelated snapshots extracted from 1 ns-long GFN2-xTB MD simulations accounting explicitly for solvent molecules using solvation spheres of increasingly large radii of 7, 9, 11, and 12 Å. Note that when considering a sphere of 10 Å, a longer 3 ns-long GFN2-xTB MD simulation was run. Convoluted spectra used a damping factor $\Gamma = 0.05$ eV. No energy shifts were applied.

experiment is striking especially when using at least a solvation sphere of 11 Å. As for 1PA, the nonhomogeneous broadening of the second absorption band featuring two peaks between 3.2 and 3.8 eV is well-captured by the ADQM-MD methodology due to the dynamic interaction of the isoalloxazine ring with its nearby surroundings. Figure S7, Supporting Information, provides more details about the influence of longer MD simulations for the 10 Å system. In 2014, List et al.[4] computed the 2PA spectrum of FMN in aqueous solution at the QM/MM level (PE-CAM-B3LYP) considering only FMN at the QM level. The average spectrum was computed for 50 snapshots of a classical MD simulation followed by QM/MM (B3LYP/6-31+G*//OPLS2005) optimizations. **Figure 24** presents this spectrum in comparison with experiment

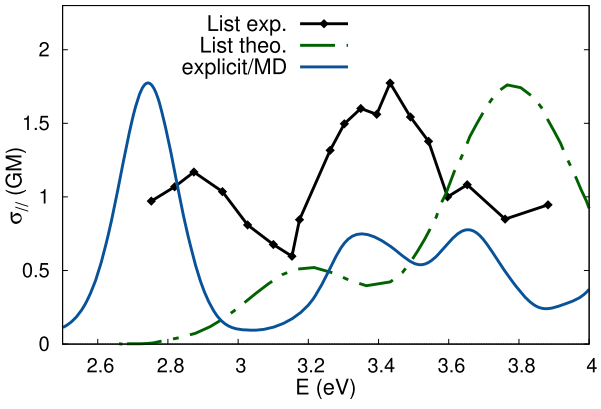

Figure 24. Experimental 2PA spectrum of FMN recorded in water by List et al.[4] in comparison with computed average dt-sTD-DFT-xTB/GBSA(water) spectrum on uncorrelated snapshots extracted from 1 ns-long GFN2-xTB MD simulations accounting explicitly for solvent molecules using a solvation sphere of 12 Å, as well as a QM-MM average spectrum computed by List et al.[4] at the PE-CAM-B3LYP/cc-pVDZ+ level on 50 snapshots of a classical MD simulation followed by QM/MM (B3LYP/6-31+G*//OPLS2005) optimizations (only FMN was accounted for into the QM layer). Convoluted spectra used a damping factor $\Gamma = 0.05$ eV (this work) and $\Gamma = 0.35$ eV (List et al.[4]). No energy shifts were applied.

and our average dt-sTD-DFT/GBSA(water) spectrum considering a solvation sphere of 12 Å. Note that the amount of explicit water molecules accounted for at the MM level by List et al.[4] is comparable to our. The QM/MM spectrum[4] is shifted by 0.4 eV with respect to experiment and does not capture the inhomogeneous broadening of the second experimental band, but reproduces the energy separation between both maxima, as well as their relative intensities. In comparison, our average dt-sTD-DFT/GBSA(water) 2PA spectrum provides peak energy positions very similar to experimental one and is able to recover the two peaks featured by the second absorption band. This result could be improved by employing other simplified schemes such as the XsTD-DFT method[91,92] as one of us did to compute 1PA and CD spectra of PYP using CAM-B3LYP.[93] Nevertheless, these results strongly advocates for the use of the ADQM-MD scheme to evalute 1- and 2PA spectra of dyes in solution.

5. Conclusions

In this study, different AQM workflows were assessed to compute 1- and 2PA of realistic systems. While neglecting dynamic structural effects but still explicitly accounting for the chromophore surroundings, the performance of ASQM scheme was evaluated for two extremely challenging systems: bR (≈3850 atoms) and iLOV (≈2000 atoms), pushing the computational boundaries of this methodology. It was also the opportunity to test with success the dt-sTD-DFT scheme to reduce the computational cost of the sTD-DFT calculations. Results showed that to reproduce experimental 1- and 2PA spectra, it is important to account for the whole protein at a QM level because absorbing amino acid residues such as tryptophan or tyrosine have non-negligible impact on the higher-energy part of the spectra (>4 eV). All spectral features were not fully captured by the ASQM scheme but still comparisons with respect to experiment were rather reasonable. Future studies are planned to further reduce the gap between simulated and real systems by including dynamic structural effects using an ADQM procedure for proteins.

Furthermore, two ADQM schemes were tested to evaluate 1- and 2 PA of FMN in aqueous solution. While the ADQM-Botlz. was not improving results with respect to the single structure approach, the ADQM-MD provided striking agreements with respect to experiment for both 1- and 2PA when including solvation spheres of at least 11 Å. In both spectra, the inhomogeneous broadening of the second absorption band results from a CT excitation modulated by the dynamic interaction of the HPO_4^- group as well as water molecules with the isoalloxazine ring. These interactions can only be coped with by an ADQM scheme. Note that the same feature is also present in iLOV spectra for which FMN is the chromophore and could be capture using such scheme. We clearly observed that explicitly accounting for both solvent and dynamic structural effects has a far more dramatic impact on spectra than simply choosing between both RI-CC2 or sTD-DFT-xTB methods in a single structure approach. Thus, these results clearly advocate for the use of an ADQM-MD workflow to compute 1- and 2PA of molecules in solution and to provide quantitative agreement with respect to experiment. Clearly,

sQC methods such as the (dt-)sTD-DFT-xTB schemes are crucial for AQM workflows. The establishment of AQM methodologies with the help of sQC methods opens the gate for a plethora of applications on realistic systems in reasonable computational times.

Acknowledgements

S.L. and M.de W. thank Prof. Dr. Stefan Grimme for his support all along these years. M.M. thanks the F.R.S.-F.N.R.S. for her Research Fellow position. Some of the computational resources were provided by the supercomputing facilities at the UCLouvain (CISM) and the Consortium des Équipements de Calcul Intensif en Fédération Wallonie Bruxelles (CÉCI) funded by the Fond de la Recherche Scientifique de Belgique (F.R.S.-FNRS) under convention 2.5020.11 and the Walloon Region.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Sarah Löffelsender: data curation (lead); formal analysis (lead); investigation (lead); methodology (equal); writing—original draft (equal); writing—review & editing (equal). **Marilù G. Maraldi**: writing—original draft (equal); writing— review & editing (supporting). **Marc de Wergifosse**: conceptualization (lead); funding acquisition (lead); methodology (equal); software (lead); supervision (lead); writing— original draft (equal); writing—review & editing (equal).

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Keywords: all-atom quantum mechanics · realistic system · simplified quantum chemistry · sTD-DFT · two-photon absorption

[1] T. J. Zuehlsdorff, C. M. Isborn, *Int. J. Quantum Chem.* **2019**, *119*, e25719.
[2] T. Giovannini, C. Cappelli, *Chem.ical Commun.ications* **2023**, *59*, 5644.
[3] T. Andruniów, M. Olivucci, in *Volume 31 Of Challenges And Advances In Computational Chemistry And Physics* Springer, New York **2021**.
[4] N. H. List, F. M. Pimenta, L. Holmegaard, R. L. Jensen, M. Etzerodt, T. Schwabe, J. Kongsted, P. R. Ogilby, O. Christiansen, *Phys. Chem. Chem. Phys.* **2014**, *16*, 9950.
[5] T. Giovannini, P. Lafiosca, C. Cappelli, *J. Chem. Theory Comput.* **2017**, *13*, 4854.
[6] T. Giovannini, F. Egidi, C. Cappelli, *Chem. Soc. Rev.* **2020**, *49*, 5664.
[7] J. M. Milanese, M. R. Provorse, E. J. Alameda, C. M. Isborn, *J. Chem. Theory Comput.* **2017**, *13*, 2159.
[8] O. Valsson, P. Campomanes, I. Tavernelli, U. Rothlisberger, C. Filippi, *J. Chem. Theory Comput.* **2013**, *9*, 2441.
[9] A. Warshel, M. Levitt, *J. Mol. Biol.* **1976**, *103*, 227.
[10] M. Svensson, S. Humbel, R. D. Froese, T. Matsubara, S. Sieber, K. Morokuma, *J. Phys. Chem.* **1996**, *100*, 19357.
[11] F. Maseras, K. Morokuma, *J. Comput. Chem.* **1995**, *16*, 1170.
[12] S. Humbel, S. Sieber, K. Morokuma, *J. Chem. Phys.* **1996**, *105*, 1959.
[13] H. M. Senn, W. Thiel, *Angew. Chem. Int. Ed.* **2009**, *48*, 1198.
[14] P. Sherwood, A. H. de Vries, M. F. Guest, G. Schreckenbach, C. R. A. Catlow, S. A. French, A. A. Sokol, S. T. Bromley, W. Thiel, A. J. Turner, S. Billeter, F. Terstegen, S. Thiel, J. Kendrick, S. C. Rogers, J. Casci, M. Watson, F. King, E. Karlsen, M. Sjøvoll, A. Fahmi, A. Schäfer, C. Lennartz, *J. Mol. Struct.: Theochem* **2003**, *632*, 1.
[15] H. Lin, D. G. Truhlar, *Theor. Chem. Acc.* **2007**, *117*, 185.
[16] L. Goletto, S. Gómez, J. H. Andersen, H. Koch, T. Giovannini, *Phys. Chem. Chem. Phys.* **2022**, *24*, 27866.
[17] C. Cappelli, *Int. J. Quantum Chem.* **2016**, *116*, 1532.
[18] C. Reichardt, T. Welton, in *Solvents and Solvent Effects in Organic Chemistry*, John Wiley & Sons, Hoboken, NJ **2011**.
[19] L. Jensen, P. T. Van Duijnen, J. G. Snijders, *J. Chem. Phys.* **2003**, *119*, 12998.
[20] C. Curutchet, A. Munoz-Losa, S. Monti, J. Kongsted, G. D. Scholes, B. Mennucci, *J. Chem. Theory Comput.* **2009**, *5*, 1838.
[21] P. Reinholdt, J. Kongsted, J. M. H. Olsen, *J. Phys. Chem. Lett.* **2017**, *8*, 5949.
[22] D. Loco, É. Polack, S. Caprasecca, L. Lagardere, F. Lipparini, J.-P. Piquemal, B. Mennucci, *J. Chem. Theory Comput.* **2016**, *12*, 3654.
[23] M. S. Gordon, L. Slipchenko, H. Li, J. H. Jensen, in *Annual Reports in Computational Chemistry*, (Ed: D. Spellmeyer, R. Wheeler) Vol3, Elsevier, Amsterdam **2007**, pp. 177–193.
[24] A. DeFusco, N. Minezawa, L. V. Slipchenko, F. Zahariev, M. S. Gordon, *J. Phys. Chem. Lett.* **2011**, *2*, 2184.
[25] L. V. Slipchenko, *J. Phys. Chem. A* **2024**, *128*, 656.
[26] C. I. Viquez Rojas, L. V. Slipchenko, *J. Chem. Theory Comput.* **2020**, *16*, 6408.
[27] D. Kosenkov, L. V. Slipchenko, *J. Phys. Chem. A* **2011**, *115*, 392.
[28] A. Severo Pereira Gomes, C. R. Jacob, *Annu. Rep. Progr. Chem. Sect. C: Phys. Chem.* **2012**, *108*, 222.
[29] E. D. Larsson, P. Reinholdt, E. D. Hedegård, J. Kongsted, *J. Phys. Chem. B* **2023**, *127*, 9905.
[30] M. Jansen, P. Reinholdt, E. D. Hedegård, C. König, *J. Phys. Chem. A* **2023**, *127*, 5689.
[31] S. Alexander, G. E. Koehl, M. Hirschberg, E. K. Geissler, P. Friedl, *Histochem. Cell Biol.* **2008**, *130*, 1147.
[32] J. Paoli, M. Smedh, M. B. Ericson, in *Seminars in Cutaneous Medicine and Surgery*, Vol. *28*, WB Saunders, Philadelphia, PA **2009**, pp. 190–195.
[33] K. Tanaka, Y. Toiyama, Y. Okugawa, M. Okigami, Y. Inoue, K. Uchida, T. Araki, Y. Mohri, A. Mizoguchi, M. Kusunoki, *Am. J. Transl. Res.* **2014**, *6*, 179.
[34] K. W. Dunn, R. M. Sandoval, K. J. Kelly, P. C. Dagher, G. A. Tanner, S. J. Atkinson, R. L. Bacallao, B. A. Molitoris, *Am. J. Physiol.-Cell Physiol.* **2002**, *283*, C905.
[35] A. Diaspro, G. Chirico, M. Collini, *Q. Rev. Biophys.* **2005**, *38*, 97.
[36] S. Ashworth, R. Sandoval, G. Tanner, B. Molitoris, *Kidney Int.* **2007**, *72*, 416.
[37] K. Svoboda, R. Yasuda, *Neuron* **2006**, *50*, 823.
[38] R. K. Benninger, D. W. Piston, *Curr. Protoc. Cell Biol.* **2013**, *59*, 4.
[39] C. Huang, J. R. Maxey, S. Sinha, J. Savall, Y. Gong, M. J. Schnitzer, *Nat. Commun.* **2018**, *9*, 872.
[40] J. Demas, J. Manley, F. Tejera, K. Barber, H. Kim, F. M. Traub, B. Chen, A. Vaziri, *Nat. Methods* **2021**, *18*, 1103.
[41] F. Helmchen, M. S. Fee, D. W. Tank, W. Denk, *Neuron* **2001**, *31*, 903.
[42] W. Zong, H. A. Obenhaus, E. R. Skytøen, H. Eneqvist, N. L. de Jong, R. Vale, M. R. Jorge, M.-B. Moser, E. I. Moser, *Cell* **2022**, *185*, 1240.
[43] T. Fischer, N. A. Hampp, *Biophys. J.* **2005**, *89*, 1175.
[44] B. Yao, Z. Ren, N. Menke, Y. Wang, Y. Zheng, M. Lei, G. Chen, N. Hampp, *Appl. Opt.* **2005**, *44*, 7344.
[45] D. Oesterhelt, C. Bräuchle, N. Hampp, *Q. Rev. Biophys.* **1991**, *24*, 425.
[46] J. Olsen, P. Jørgensen, *J. Chem. Phys.* **1985**, *82*, 3235.
[47] T. Helgaker, S. Coriani, P. Jørgensen, K. Kristensen, J. Olsen, K. Ruud, *Chem. Rev.* **2012**, *112*, 543.
[48] M. J. Paterson, O. Christiansen, F. Pawłowski, P. Jørgensen, C. Hättig, T. Helgaker, P. Sałek, *J. Chem. Phys.* **2006**, *124*, 054322.
[49] K. D. Nanda, A. I. Krylov, *J. Chem. Phys.* **2015**, 142.
[50] C. Hättig, O. Christiansen, P. Jørgensen, *J. Chem. Phys.* **1998**, *108*, 8355.
[51] D. H. Friese, C. Hättig, K. Ruud, *Phys. Chem. Chem. Phys.* **2012**, *14*, 1175.
[52] M. A. Salem, A. Brown, *J. Chem. Theory Comput.* **2014**, *10*, 3260.
[53] M. T. P. Beerepoot, D. H. Friese, K. Ruud, *Phys. Chem. Chem. Phys.* **2014**, *16*, 5958.
[54] M. T. Beerepoot, D. H. Friese, N. H. List, J. Kongsted, K. Ruud, *Phys. Chem. Chem. Phys.* **2015**, *17*, 19306.
[55] M. T. P. Beerepoot, M. M. Alam, J. Bednarska, W. Bartkowiak, K. Ruud, R. Zaleśny, *J. Chem. Theory Comput.* **2018**, *14*, 3677.
[56] I. A. Elayan, A. Brown, *J. Phys. Chem. A* **2024**, *128*, 7511.

[57] M. Chołuj, M. M. Alam, M. T. P. Beerepoot, S. P. Sitkiewicz, E. Matito, K. Ruud, R. Zaleśny, *J. Chem. Theory Comput.* **2022**, *18*, 1046.
[58] K. Ahmadzadeh, X. Li, Z. Rinkevicius, P. Norman, R. Zaleśny, *J. Phys. Chem. Lett.* **2024**, *15*, 969.
[59] C. Naim, R. Zaleśny, D. Jacquemin, *J. Chem. Theory Comput.* **2024**, *20*, 9093.
[60] P. Sałek, O. Vahtras, J. Guo, Y. Luo, T. Helgaker, H. Ågren, *Chem. Phys. Lett.* **2003**, *374*, 446.
[61] F. Zahariev, M. S. Gordon, *J. Chem. Phys.* **2014**, *140*, 18A523.
[62] S. M. Parker, D. Rappoport, F. Furche, *J. Chem. Theory Comput.* **2018**, *14*, 807.
[63] M. Paterson, J. Kongsted, O. Christiansen, K. Mikkelsen, C. Nielsen, *J. Chem. Phys.* **2006**, *125*, 184501.
[64] D. L. Silva, N. A. Murugan, J. Kongsted, Z. Rinkevicius, S. Canuto, H. Ågren, *J. Phys. Chem. B* **2012**, *116*, 8169.
[65] J. Olesiak-Banska, K. Matczyszyn, R. Zaleśny, N. A. Murugan, J. Kongsted, H. Ågren, W. Bartkowiak, M. Samoc, *J. Phys. Chem. B* **2013**, *117*, 12013.
[66] K. Matczyszyn, J. Olesiak-Banska, K. Nakatani, P. Yu, N. A. Murugan, R. Zaleśny, A. Roztoczyńska, J. Bednarska, W. Bartkowiak, J. Kongsted, H. Ågren, M. Samoć, *J. Phys. Chem. B* **2015**, *119*, 1515.
[67] T. N. Ramos, L. R. Franco, D. L. Silva, S. Canuto, *J. Chem. Phys.* **2023**, *159*, 024309.
[68] M. Wielgus, R. Zaleśny, N. A. Murugan, J. Kongsted, H. Ågren, M. Samoc, W. Bartkowiak, *ChemPhysChem* **2013**, *14*, 3731.
[69] N. Arul Murugan, J. Kongsted, Z. Rinkevicius, K. Aidas, K. V. Mikkelsen, H. Ågren, *Phys. Chem. Chem. Phys.* **2011**, *13*, 12506.
[70] A. H. Steindal, J. M. H. Olsen, K. Ruud, L. Frediani, J. Kongsted, *Phys. Chem. Chem. Phys.* **2012**, *14*, 5440.
[71] N. A. Murugan, R. Zaleśny, *J. Chem. Inf. Model.* **2020**, *60*, 3854.
[72] M. Rossano-Tapia, J. M. H. Olsen, A. Brown, *Front. Mol. Biosci.* **2020**, *7*.
[73] K. Sneskov, J. M. H. Olsen, T. Schwabe, C. Hättig, O. Christiansen, J. Kongsted, *Phys. Chem. Chem. Phys.* **2013**, *15*, 7567.
[74] A. A. A. Attia, P. Reinholdt, J. Kongsted, *Adv. Theory Simul.* **2021**, *4*, 2000294.
[75] J. Olsen, N. List, K. Kristensen, J. Kongsted, *J. Chem. Theory Comput.* **2015**, *11*, 1832.
[76] K. Schwinn, N. Ferré, M. Huix-Rotllant, *Phys. Chem. Chem. Phys.* **2020**, *22*, 12447.
[77] M. P. Kabir, Y. Orozco-Gonzalez, S. Gozem, *Phys. Chem. Chem. Phys.* **2019**, *21*, 16526.
[78] R. K. Kar, A.-F. Miller, M.-A. Mroginski, *WIREs Comput. Mol. Sci.* **2022**, *12*, e1541.
[79] G. Palczewska, F. Vinberg, P. Stremplewski, M. P. Bircher, D. Salom, K. Komar, J. Zhang, M. Cascella, M. Wojtkowski, V. J. Kefalov, K. Palczewski, *Proc. Natl. Acad. Sci.* **2014**, *111*, E5445.
[80] S. Grimme, *J. Chem. Phys.* **2013**, *138*, 244104.
[81] C. Bannwarth, S. Grimme, *Comput. Theor. Chem.* **2014**, *1040–1041*, 45.
[82] S. Grimme, C. Bannwarth, *J. Chem. Phys.* **2016**, *145*, 054103.
[83] M. de Wergifosse, S. Grimme, *J. Chem. Phys.* **2018**, *149*, 024108.
[84] M. de Wergifosse, S. Grimme, *J. Chem. Phys.* **2019**, *150*, 094112.
[85] M. de Wergifosse, C. Bannwarth, S. Grimme, *J. Phys. Chem. A* **2019**, *123*, 5815.
[86] M. de Wergifosse, J. Seibert, S. Grimme, *J. Chem. Phys.* **2020**, *153*, 084116.
[87] J. Seibert, B. Champagne, S. Grimme, M. de Wergifosse, *J. Chem. Phys. B* **2020**, *124*, 2568.
[88] P. Beaujean, B. Champagne, S. Grimme, M. de Wergifosse, *J. Chem. Phys. Lett.* **2021**, *12*, 9684.
[89] M. de Wergifosse, S. Grimme, *J. Phys. Chem. A* **2021**, *125*, 3841.
[90] M. de Wergifosse, P. Beaujean, S. Grimme, *J. Phys. Chem. A* **2022**, *126*, 7534.
[91] S. Löffelsender, P. Beaujean, M. de Wergifosse, *WIREs Comput. Mol. Sci.* **2024**, *14*, e1695.
[92] M. de Wergifosse, S. Grimme, *J. Chem. Phys.* **2024**, *160*, 204110.
[93] M. de Wergifosse, *J. Phys. Chem. Lett.* **2024**, *15*, 12628.
[94] D. Oesterhelt, W. Stoeckenius, *Nat. New Biol.* **1971**, *233*, 149.
[95] U. Haupts, J. Tittor, D. Oesterhelt, *Annu. Rev. Biophys.* **1999**, *28*, 367.
[96] M. P. Krebs, T. A. Isenbarger, *Biochim. et Biophys. Acta Bioenergetics* **2000**, *1460*, 15.
[97] J. K. Lanyi, *Annu. Rev. Physiol.* **2004**, *66*, 665.
[98] K. Hayashi, M. Tanigawara, J. Kishikawa, *Biophysics* **2012**, *8*, 67.
[99] K. Yanai, K. Ishimura, A. Nakayama, J. Hasegawa, *J. Chem. Theory Comput.* **2018**, *14*, 3643.
[100] G. N. Kovács, J.-P. Colletier, M. L. Grünbein, Y. Yang, T. Stensitzki, A. Batyuk, S. Carbajo, R. B. Doak, D. Ehrenberg, L. Foucar, R. Gasper, A. Gorel, M. Hilpert, M. Kloos, J. E. Koglin, J. Reinstein, C. M. Roome, R. Schlesinger, M. H. Seaberg, R. L. Shoeman, M. Stricker, S. Boutet, S. Haacke, J. Heberle, K. Heyne, T. Domratcheva, T. R. M. Barends, I. Schlichting, *Nat. Commun.* **2019**, *10*, 3177.
[101] S. Hayashi, E. Tajkhorshid, E. Pebay-Peyroula, A. Royant, E. M. Landau, J. Navarro, K. Schulten, *J. Phys. Chem. B* **2001**, *105*, 10124.
[102] R. R. Birge, C.-F. Zhang, *J. Chem. Phys.* **1990**, *92*, 7178.
[103] M. B. Masthay, D. M. Sammeth, M. C. Helvenston, C. B. Buckman, W. Li, M. J. Cde-Baca, J. T. Kofron, *J. Am. Chem. Soc.* **2002**, *124*, 3418.
[104] D. Rhinow, M. Imhof, I. Chizhik, R.-P. Baumann, N. Hampp, *J. Phys. Chem. B* **2012**, *116*, 7455.
[105] S. Chapman, C. Faulkner, E. Kaiserli, C. Garcia-Mata, E. I. Savenkov, A. G. Roberts, K. J. Oparka, J. M. Christie, *Proc. Natl. Acad. Sci.* **2008**, *105*, 20038.
[106] L. M. Antill, S.-Y. Takizawa, S. Murata, J. R. Woodward, *Mol. Phys.* **2019**, *117*, 2594.
[107] S. S. Daz, H. Al-Zubaidi, A. C. Ross-Obare, S. O. Obare, *Phys. Sci. Rev.* **2020**, *5*, 20200007.
[108] X. Ran, Q. Zhang, Y. Zhang, J. Chen, Z. Wei, Y. He, L. Guo, *Molecules* **2020**, *25*, 2544.
[109] J. Torra, C. Lafaye, L. Signor, S. Aumonier, C. Flors, X. Shu, S. Nonell, G. Gotthard, A. Royant, *Sci. Rep.* **2019**, *9*, 2428.
[110] S. Salzmann, M. R. Silva-Junior, W. Thiel, C. M. Marian, *J. Phys. Chem. B* **2009**, *113*, 15610.
[111] C. Neiss, P. Saalfrank, M. Parac, S. Grimme, *J. Phys. Chem. A* **2003**, *107*, 140.
[112] R. J. Stanley, H. Jang, *J. Phys. Chem. A* **1999**, *103*, 8976.
[113] M. P. Kabir, P. Ghosh, S. Gozem, *J. Phys. Chem. B* **2024**, *128*, 7545.
[114] A. Onufriev, D. Bashford, D. A. Case, *Proteins: Struct. Funct. Bioinf.* **2004**, *55*, 383.
[115] G. Sigalov, A. Fenley, A. Onufriev, *J. Chem. Phys.* **2006**, *124*, 124902.
[116] A. W. Lange, J. M. Herbert, *J. Chem. Theory Comput.* **2012**, *8*, 1999.
[117] Y. De Coene, S. Van Cleuvenbergen, N. Van Steerteghem, V. Baekelandt, T. Verbiest, C. Bartic, K. Clays, *J. Phys. Chem. C* **2017**, *121*, 6909.
[118] E. Winder, M. H. Griep, D. Lueking, C. Friedrich, in *Proc. of the 26th Army Science Conf.*, Orlando, Florida, Vol. *14*, **2008**.
[119] R. J. Homans, R. U. Khan, M. B. Andrews, A. E. Kjeldsen, L. S. Natrajan, S. Marsden, E. A. McKenzie, J. M. Christie, A. R. Jones, *Phys. Chem. Chem. Phys.* **2018**, *20*, 16949.
[120] P. Pracht, F. Bohle, S. Grimme, *Phys. Chem. Chem. Phys.* **2020**, *22*, 7169.
[121] J. W. Furness, A. D. Kaplan, J. Ning, J. P. Perdew, J. Sun, *J. Phys. Chem. Lett.* **2020**, *11*, 8208.
[122] S. Grimme, A. Hansen, S. Ehlert, J.-M. Mewes, *J. Chem. Phys.* **2021**, *154*, 064103.
[123] E. Caldeweyher, C. Bannwarth, S. Grimme, *J. Chem. Phys.* **2017**, *147*, 034112.
[124] S. Grimme, F. Bohle, A. Hansen, P. Pracht, S. Spicher, M. Stahn, *J. Phys. Chem. A* **2021**, *125*, 4039.
[125] Universität Bonn, M. C. T. C., Bonn, Germany, F. Bohle, CENSO - Commandline Energetic Sorting of Conformer-Rotamer Ensembles, Version 1.2.0, https://github.com/grimme-lab/censo/releases/, (acessed: March 2022).
[126] V. Ásgeirsson, C. A. Bauer, S. Grimme, *Chem. Sci.* **2017**, *8*, 4879.
[127] T. J. Zuehlsdorff, A. Montoya-Castillo, J. A. Napoli, T. E. Markland, C. M. Isborn, *J. Chem. Phys.* **2019**, *151*, 074111.
[128] J. Cerezo, D. Aranda, F. J. Avila Ferrer, G. Prampolini, F. Santoro, *J. Chem. Theory Comput.* **2020**, *16*, 1215.
[129] R. Crespo-Otero, M. Barbatti, *Theor. Chem. Acc.* **2012**, *131*, 1237.
[130] V. Lukeš, R. Šolc, M. Barbatti, H. Lischka, H.-F. Kauffmann, *J. Theor. Computat. Chem.* **2010**, *09*, 249.
[131] P. Löwdin, *J. Chem. Phys.* **1950**, *18*, 365.
[132] K. Nishimoto, N. Mataga, *Z. Phys. Chem* **1957**, *12*, 335.
[133] K. Ohno, *Theor. Chim. Acta* **1964**, *2*, 219.
[134] G. Klopman, *J. Am. Chem. Soc.* **1964**, *86*, 4550.
[135] R. J. Buenker, S. D. Peyerimhoff, *Theor. Chim. Acta* **1974**, *35*, 33.
[136] R. J. Buenker, S. D. Peyerimhoff, *Theor. Chim. Acta* **1975**, *39*, 217.
[137] S. Grimme, M. Waletzke, *J. Chem. Phys.* **1999**, *111*, 5645.
[138] F. Neese, *J. Chem. Phys.* **2003**, *119*, 9428.
[139] P. Monson, W. McClain, *J. Chem. Phys.* **1970**, *53*, 29.
[140] M. J. Wirth, A. Koskelo, M. J. Sanders, *Appl. Spectrosc.* **1981**, *35*, 14.
[141] S. Chakrabarti, K. Ruud, *J. Phys. Chem. A* **2009**, *113*, 5485.
[142] R. Nifos, Y. Luo, *J. Phys. Chem. B* **2007**, *111*, 14043.
[143] D. H. Friese, M. T. Beerepoot, M. Ringholm, K. Ruud, *J. Chem. Theory Comput.* **2015**, *11*, 1129.
[144] S. Gholami, L. Pedraza-González, X. Yang, A. A. Granovsky, I. N. Ioffe, M. Olivucci, *The J. Phys. Chem. Lett.* **2019**, *10*, 6293.

[145] A. Rizzo, S. Coriani, K. Ruud, in *Response Function Theory Computational Approaches To Linear And Nonlinear Optical Spectroscopy*, John Wiley and Sons, Hoboken, NJ **2012**, pp. 77–135.
[146] M. de Wergifosse, A. L. Houk, A. I. Krylov, C. G. Elles, *J. Chem. Phys.* **2017**, *146*, 144305.
[147] J.-D. Chai, M. Head-Gordon, *J. Chem. Phys.* **2008**, *128*, 084106.
[148] J.-D. Chai, M. Head-Gordon, *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615.
[149] R. Ditchfield, W. J. Hehre, J. A. Pople, *J. Chem. Phys.* **1971**, *54*, 724.
[150] M. M. Francl, W. J. Pietro, W. J. Hehre, J. S. Binkley, M. S. Gordon, D. J. DeFrees, J. A. Pople, *J. Chem. Phys.* **1982**, *77*, 3654.
[151] M. S. Gordon, J. S. Binkley, J. A. Pople, W. J. Pietro, W. J. Hehre, *J. Am. Chem. Soc.* **1982**, *104*, 2797.
[152] P. C. Hariharan, J. A. Pople, *Theor. Chim. Acta* **1973**, *28*, 213.
[153] W. J. Hehre, R. Ditchfield, J. A. Pople, *J. Chem. Phys.* **1972**, *56*, 2257.
[154] B. P. Pritchard, D. Altarawy, B. Didier, T. D. Gibsom, T. L. Windus, *J. Chem. Inf. Model.* **2019**, *59*, 4814.
[155] C. Bannwarth, S. Ehlert, S. Grimme, *J. Chem. Theory Comput.* **2019**, *15*, 1652.
[156] S. Grimme, *J. Comput. Chem.* **2006**, *27*, 1787.
[157] S. Grimme, S. Ehrlich, L. Goerigk, *J. Comput. Chem.* **2011**, *32*, 1456.
[158] F. Weigend, R. Ahlrichs, *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297.
[159] S. Sinnecker, A. Rajendran, A. Klamt, M. Diedenhofen, F. Neese, *J. Phys. Chem. A* **2006**, *110*, 2235.
[160] M. Renz, M. Kess, M. Diedenhofen, A. Klamt, M. Kaupp, *J. Chem. Theory Comput.* **2012**, *8*, 4189.
[161] S. Grimme, *Chemistry* **2012**, *18*, 9955.
[162] S. Spicher, S. Grimme, *J. Phys. Chem. Lett.* **2020**, *11*, 6606.
[163] S. Spicher, S. Grimme, *J. Chem. Theory Comput.* **2021**, *17*, 1701.
[164] A. Klamt, *J. Phys. Chem.* **1995**, *99*, 2224.
[165] A. Klamt, V. Jonas, T. Bürger, J. C. W. Lohrenz, *J. Phys. Chem. A* **1998**, *102*, 5074.
[166] S. G. Balasubramani, G. P. Chen, S. Coriani, M. Diedenhofen, M. S. Frank, Y. J. Franzke, F. Furche, R. Grotjahn, M. E. Harding, C. Hättig, A. Hellweg, B. Helmich-Paris, C. Holzer, U. Huniar, M. Kaupp, A. Marefat Khah, S. Karbalaei Khani, T. Müller, F. Mack, B. D. Nguyen, S. M. Parker, E. Perlt, D. Rappoport, K. Reiter, S. Roy, M. Rückert, G. Schmitz, M. Sierka, E. Tapavicza, D. P. Tew, C. van Wüllen, V. K. Voora, F. Weigend, A. Wodyński, J. M. Yu, *J. Chem. Phys.* **2020**, *152*, 184107.
[167] F. Furche, R. Ahlrichs, C. Hättig, W. Klopper, M. Sierka, F. Weigend, *WIREs Computat. Mol. Sci.* **2014**, *4*, 91.
[168] Universität Karlsruhe & Forschungszentrum Karlsruhe GmbH, Karlsruhe, Germany, TURBOMOLE 7.6, https://www.turbomole.org/, (accessed: March 2022).
[169] F. Eckert, A. Klamt, *AIChE J.* **2002**, *48*, 369.
[170] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, *J. Comput. Chem.* **2004**, *25*, 1605.
[171] M. Brehm, B. Kirchner, *J. Chem. Inf. Model.* **2011**, *51*, 2007.
[172] M. Brehm, M. Thomas, S. Gehrke, B. Kirchner, *J. Chem. Phys.* **2020**, *152*, 164105.
[173] C. Hättig, F. Weigend, *J. Chem. Phys.* **2000**, *113*, 5154.
[174] J. Dunning, H. Thom, *J. Chem. Phys.* **1989**, *90*, 1007.
[175] Universität Karlsruhe & Forschungszentrum Karlsruhe GmbH, Karlsruhe, Germany, TURBOMOLE 7.7.1, https://www.turbomole.org/, (accessed: October 2023).
[176] C. Bannwarth, S. Grimme, sTDA-xTB for Ground State Calculations, https://github.com/grimme-lab/xtb4stda, **2016**.
[177] std2, A Program for Computing Excited States and Response Functions Via Simplified TD-DFT methods (sTDA, sTD-DFT, SF-sTD-DFT, XsTDA, XsTD-DFT, and SF-Xs-TD-DFT), https://github.com/grimme-lab/std2, (accessd: January **2025**).
[178] N. Nunthaboot, A. Nueangaudom, K. Lugsanangarm, S. Pianwanit, S. Kokpol, F. Tanaka, *Comput. Biol. Chem.* **2018**, *72*, 96.

Manuscript received: December 16, 2024
Revised manuscript received: April 7, 2025
Version of record online: