Initial Maximum Overlap Method for Large Systems through the QM/ELMO Embedding Technique

Giovanni Macetti[(1)], Alessandro Genoni[(1)*]

(1) Université de Lorraine & CNRS, Laboratoire de Physique et Chimie Théoriques (LPCT), UMR CNRS 7019, 1 Boulevard Arago, F-57078 Metz, France.

* Correspondence to:

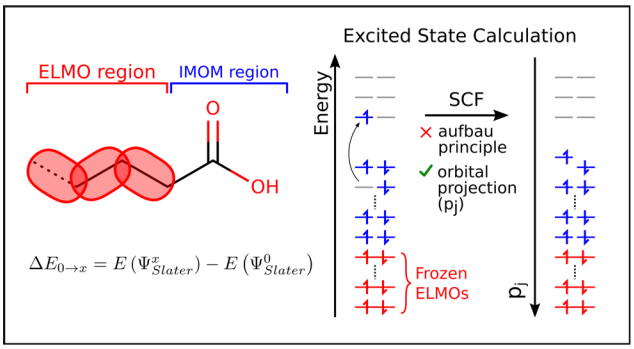
- Alessandro Genoni, Université de Lorraine & CNRS, Laboratoire de Physique et Chimie Théoriques (LPCT), UMR CNRS 7019, 1 Boulevard Arago, F-57078 Metz, France. E-mail: alessandro.genoni@univ-lorraine.fr; Phone: +33 (0)3 72 74 91 70; Fax: +33 (0)3 72 74 91 87.

ABSTRACT

Quantum chemistry offers a large variety of methods to treat excited states. Many of them are based on a multi-reference wave function *ansatz* and are therefore characterized by an intrinsic complexity and high computational costs. To overcome these drawbacks and also some limitations of simpler single-reference approaches (e.g., Configuration Interaction Singles (CIS) and Time-Dependent Density Functional Theory (TDDFT)), the single-determinant ΔSelf-Consistent Field-Initial Maximum Overlap Method (ΔSCF-IMOM) has been proposed. This strategy substitutes the *aufbau principle* with a criterion that occupies molecular orbitals at successive SCF iterations on the basis of their maximum overlap with a proper set of guess orbitals for the target excited state. In this way it prevents the SCF process to collapse to the ground state wave function and provides excited state single Slater determinant solutions to the SCF equations. Here we propose to extend the applicability of IMOM to the treatment of localized excited states of large systems. To accomplish this task, we coupled it with the QM/ELMO (quantum mechanics / extremely localized molecular orbitals) strategy, a quantum mechanical embedding method in which the most chemically relevant part of the system is treated through traditional quantum chemical approaches, while the rest is described by extremely localized molecular orbitals transferred from recently constructed libraries or proper model molecules. After presenting the theoretical foundations of the new IMOM/ELMO technique, in this paper we will show and discuss the results of preliminary test calculations carried out on both model systems (i.e., decanoic acid, decene, decapentaene and solvated acrolein) and a system of biological interest (flavin mononucleotide in the flavodoxin protein). We observed that, for localized excited states, the new IMOM/ELMO method provides reliable results, and it reproduces the

outcomes of fully IMOM calculations within the chemical accuracy threshold (i.e., 0.043 eV) by including only a limited number of atoms in the QM region. Furthermore, the first application of our embedding technique to a larger biological system gave completely plausible results in line with those obtained through more traditional quantum mechanical methods, thus opening the possibility of using the new approach in future investigations of photobiology problems.

TOC GRAPHICS



KEYWORDS: Excited states, Initial Maximum Overlap Method (IMOM), Embedding techniques, QM/QM' strategies, extremely localized molecular orbitals (ELMOs).

I. INTRODUCTION

Excited states play a crucial role in many natural phenomena and their full understanding is of paramount importance to get fundamental insights and advancements in many research fields, ranging from materials science to biology. To achieve this goal, quantum chemistry currently offers a wide arsenal of advanced and sophisticated methods, among which we can mention techniques as Multi Reference Configuration Interaction (MRCI),[1,2] Complete Active Space Self-Consistent Field (CASSCF),[3, 4] Complete Active Space Perturbation Theory (e.g., CASPT2),[5-7] Multi-Reference Møller-Plesset (MRMP),[8] Equation-of-Motion Coupled Cluster (EOM-CC),[9-11] Linear-Response Coupled Cluster,[12,13] Configuration Interaction Singles (CIS)[14-16] and Time-Dependent Density Functional Theory (TDDFT).[17-19]

Many of the above-mentioned quantum chemistry strategies for the description of excited states are based on a multi-reference wave function *ansatz* and are consequently cumbersome and computationally expensive. These drawbacks are partially solved by the single-reference CIS and TDDFT techniques, which, on the other hand, show other non-negligible weaknesses, mainly due to the complete lack of molecular orbital relaxation in the excited states. For example, this leads to overestimations of excitation energies in the CIS case[20] and to the well-known unsatisfactory performances of the TDDFT approach in describing Rydberg[21,22] and charge-transfer[23,24] states. In the latter cases, the choice of a suitable range-separated functional usually solves the problem but may sometimes represent a complication. To avoid it, one could use the alternative Bethe-Salpeter equation (BSE) formalism[25] that is as computationally expensive as TDDFT but does not require the selection of functionals for the calculations.

However, to overcome some of the shortcomings of the above-mentioned single-references techniques and, at the same time, to keep their computational advantages and easiness, it is also possible to resort to ΔSelf-Consistent Field (ΔSCF) approaches[26-28] in which the excitation energy is obtained as difference between the ground and excited state energies at Hartree-Fock (HF) or density functional theory (DFT) level. In this context, to obtain excited state SCF solutions different from the ground state ones (namely, to avoid the so-called variational collapse), Gill and coworkers proposed a simple algorithm/strategy: the Maximum Overlap Method (MOM),[29-31] which afterwards evolved in its more stable version, currently known as Initial Maximum Overlap Method (IMOM)[32,33]. For the computation of the excited state wave function, MOM and IMOM start from a set of guess/reference orbitals that lie in the basin of attraction of the desired excited state and that are generally selected by promoting one or more electrons from the occupied to the virtual ground state molecular orbitals (MOs) according to chemical intuition. These orbitals afterwards relax during the subsequent SCF procedure, in which, at each step, the MOs are not occupied according to the *aufbau* principle but adopting a criterion that chooses those orbitals that have the largest projection onto the space of the MOs at the previous iteration (MOM case) or onto the space of the starting guess orbitals (IMOM case). As of today, MOM and IMOM are probably the simplest algorithms to obtain *non-aufbau* solutions to the SCF equations. However, it is worth pointing out that the former cannot always prevent the variational collapse due to drifts of the reference molecular orbitals away from the target *non-aufbau* state, while the latter generally avoids the previous problem but sometimes shows convergence difficulties.[34] To overcome these drawbacks, alternative algorithms have been recently proposed. Among the most remarkable ones we can mention the strategy proposed by Levi and

coworkers, who proposed a direct optimization approach combined with MOM,[35,36] the square gradient minimization (SGM) introduced by Hait and Head-Gordon,[37] and the State-Targeted Energy Projection (STEP) devised by Carter-Fenk and Herbert[38]. Finally, for the sake of completeness, it is worth reminding that ΔSCF methods are generally affected by spin-contamination because singlet open-shell excited states cannot be properly described through single Slater determinant wave functions.[34] To solve this problem, the simplest possibility consists in applying a spin purification *a posteriori*[34] through the approximate spin-projection protocol[39]. Another option is to resort to an orbital-optimized technique for excited states that is similar but different from the ΔSCF strategies: the restricted open-shell Kohn-Sham (ROKS) approach,[40-42] where the spin-corrected energy expression associated with a two-determinant (i.e., the singlet and triplet Slater determinants) wave function for an open-shell singlet excited state is minimized with respect to a single set of molecular orbitals.

Anyway, although it is important to bear in mind all the above-mentioned precisions, MOM and IMOM remain the simplest algorithms within the context of ΔSCF methods. For this reason, they will be the main subject of the present paper. In particular, we will focus on IMOM that converges to the desired excited states better than MOM.[32]

As one can imagine, the Initial Maximum Overlap Method has a scaling ($\sim M^3$, with M as the number of basis functions used in the computation) that is much more favorable than those of more sophisticated multi-reference approaches for excited states. Nevertheless, as the size of the investigated systems becomes larger, the computational cost increases also for IMOM and, more importantly, since molecular orbitals are more and more delocalized, the initial selection of the proper guess/reference orbitals for the determination of the excited state wave function

becomes less and less trivial and convergence problems may also arise. These factors prevent a straightforward application of the Initial Maximum Overlap Method to large systems in all the situations.

To extend its range of applicability, IMOM could be coupled with multi-scale embedding strategies of quantum chemistry,[43] namely techniques in which the most chemically relevant region of the system under exam is described by means of a high-level quantum mechanical approach, while the remaining part is treated through a lower-level method. Among these techniques, a prominent place is certainly occupied by the quantum mechanics / molecular mechanics (QM/MM) strategies,[44-48] which historically represent the first examples of multi-scale embedding approaches. However, in more recent years, the development of fully quantum mechanical embedding methods has also witnessed a large and rapid expansion, with the introduction of the density matrix[49-52] and density functional embedding strategies[53-68]. Some of these approaches have been also coupled with traditional quantum chemical techniques for the treatment of excited states,[69-77] and, in all these cases, the range of applicability of the parent quantum chemistry method for excited states increased without affecting the accuracy of the results.

Another fully quantum mechanical embedding approach is the recent QM/ELMO (quantum mechanics / extremely localized molecular orbital) method, which will play a crucial role in this work.[78-82] This technique derives from the Local Self-Consistent Field (LSCF) technique[83,84] and consists in treating the most important region of the investigated system through a traditional quantum chemical strategy, while the rest is described by means of extremely localized molecular orbitals (ELMOs)[85-87] properly transferred from recently constructed libraries[88-90] or suitable tailor-made model molecules. In fact, since ELMOs are molecular orbitals strictly localized on small

fragments (such as atoms, bonds or functional groups), they can be easily exported from a molecule to another.[88,89,91-95] Databanks of these orbitals have been thus assembled in order to instantaneously reconstruct approximate wave functions and electron densities of polypeptides and large proteins,[90] but also to refine crystal structures of macromolecules and organometallic compounds[96] in the framework of modern quantum crystallography[97-102] (more details about the ELMOs transferability and the ELMO libraries are given in the Supporting Information).

Recent validation studies have shown that the new QM/ELMO embedding technique is able to reproduce the corresponding fully quantum mechanical calculations within chemical accuracy at a significantly reduced computational cost.[79,80] On top of that, and interestingly for this work, very recently the QM/ELMO approach has been also successfully interfaced with Time-Dependent Density Functional Theory and Equation-of-Motion Coupled Cluster.[80] For these reasons, here we propose the coupling of IMOM with the QM/ELMO strategy in order to easily extend the applicability of the ΔSCF Initial Maximum Overlap Method. In principle, this should allow the treatment of localized excited states for large systems without impacting on the accuracy of the results.

To prove the soundness of the new IMOM/ELMO approach, the results of fully IMOM calculations (e.g., excitation energies and oscillator strengths) on relatively large systems will be used as references and will be compared to those obtained through corresponding IMOM/ELMO computations in which the size of the quantum mechanical region is gradually increased. Finally, a first application of the new technique to a problem of biological interest will be presented and discussed to show usefulness and potentialities of the new method.

II. THEORY

In this section we will briefly review the main theoretical features of the IMOM and QM/ELMO methods, which are the two strategies at the basis of the IMOM/ELMO approach proposed in this work.

II.A Initial Maximum Overlap Method (IMOM). As already mentioned in the Introduction, IMOM[32] is a ΔSCF technique that computes excitation energies as differences between the energies of the single Slater determinants associated with the ground and excited states. The method can be summarized as follows:

1. Ground-state calculation at Hartree-Fock or DFT level. This provides occupied and virtual orbitals among which it will be possible to select the set of guess/reference orbitals for the subsequent calculation of the excited state wave function (see point 3 below).
2. Selection of the guess/reference molecular orbitals for the excited state wave function computation. This is generally done by promoting one or more electrons from the set of occupied ground state MOs to the set of virtual ground state MOs, according to the excited state that one wants to target in the computation.
3. Excited state calculation at Hartree-Fock or DFT level using the MOs selected at point 2 as guess/reference orbitals. During the SCF cycle, the excited state molecular orbitals are allowed to freely relax. At each iteration, the occupation of the molecular orbitals is not decided according to the *aufbau* principle, but through a criterion that considers the projection of the current MOs onto the above-mentioned guess/reference orbitals. In other words, instead of occupying the molecular orbitals having the lowest energy, we occupy the MOs characterized by the largest projection p_j onto the guess/reference orbitals, which is given by the following equation:

$$p_j = \left[\sum_{i=1}^{N_{ref}} \left(O_{ij}\right)^2 \right]^{1/2} \qquad (1)$$

where N_{ref} is the number of guess/reference orbitals and O_{ij} is the overlap integral between the new j-th molecular orbital at the current iteration and the i-th guess/reference molecular orbital, namely

$$O_{ij} = \sum_{\mu=1}^{M} \sum_{\nu=1}^{M} C_{\mu i}^{ref} \, S_{\mu\nu} \, C_{\nu j}^{new} \qquad (2)$$

with M as the total number of basis functions used in the calculations and $S_{\mu\nu}$ as the overlap between the μ-th and ν-th basis functions

4. The excitation energy is computed as difference between the energy resulting from the excited state calculation (see point 3) and the one obtained through the corresponding ground state computation (see point 1). The two calculations are performed on the same geometry and no zero-point vibration energy correction is introduced. Therefore, the computed excitation energies are actually vertical excitation energies.

To perform all the calculations of the present work, we implemented the IMOM algorithm by adapting the SCF routines of our modified version of the *Gaussian09* quantum chemistry package.[103] Finally, it is worth noting that the molecular orbitals obtained from excited state calculations can be also exploited in post-HF computations to determine the correlation energy of the excited states. This was done at MP2 level for the original Maximum Overlap Method,[29] but this option was not considered in the present work, where we limited ourselves to IMOM calculations at DFT level (see Computational Details sections).

II.B The QM/ELMO technique. The quantum mechanics / extremely localized molecular orbital approach is a fully quantum mechanical embedding method that allows the treatment of the most important part of an examined system at a desired quantum chemical level of theory (e.g., at Hartree-Fock, DFT, post-Hartree-Fock, TDDFT, EOM-CCSD level), while the rest is described by means of transferred and frozen ELMOs.[78-82]

After defining the QM and ELMO subunits and after transferring the required extremely localized molecular orbitals to the ELMO region, the QM/ELMO procedure can be seen as consisting in two main parts: i) preliminary orthogonalization of molecular orbitals and basis functions for the real computation; ii) real QM/ELMO self-consistent field algorithm.

If $\boldsymbol{\chi} = [|\chi_1\rangle, |\chi_2\rangle, \dots , |\chi_M\rangle]$ is the $1 \times M$ array containing the M original non-orthogonal basis functions of the full system (which comprises the QM and ELMO regions) and $\boldsymbol{\chi}' = [|\chi'_1\rangle, |\chi'_2\rangle, \dots, |\chi'_{M_{QM}}\rangle]$ is the final $1 \times M_{QM}$ array containing the M_{QM} final orthonormal basis functions for the QM subunit (with $M_{QM} \ll M$), it is possible to show that the above-mentioned preliminary orthogonalization procedure can be summarized by the following transformation:

$$\boldsymbol{\chi}' = \boldsymbol{\chi}\,\mathbf{B} \quad (3)$$

where **B** is a global $M \times M_{QM}$ transformation matrix that plays a pivotal role in the QM/ELMO SCF algorithm described below (full details about the orthogonalization procedure can be found in the Supporting Information or in the original papers of the QM/ELMO approach[78,79]).

After the preliminary orthogonalization, the real QM/ELMO SCF algorithm can start. It entails the following six steps:

1. Construction of the Fock matrix **F** in the set of the original basis functions $\boldsymbol{\chi}$.

2. Transformation of the Fock matrix to the new set of orthogonal basis functions $\boldsymbol{\chi}'$ for the QM subsystem through the transformation matrix $\mathbf{B}$: $\mathbf{F}' = \mathbf{B}^{\dagger}\mathbf{F}\,\mathbf{B}$.
3. Diagonalization of the Fock matrix $\mathbf{F}'$: $\mathbf{F}'\mathbf{C}' = \mathbf{C}'\,\mathbf{E}'$.
4. Transformation of the molecular orbitals to the set of the original basis functions: $\mathbf{C} = \mathbf{B}\,\mathbf{C}'$.
5. Computation of the QM one-electron density matrix: $\mathbf{P}^{QM} = \mathbf{C}\mathbf{C}^{\dagger}$.
6. Check of convergence on energy and density matrix: if convergence is achieved, the SCF procedure ends; otherwise the cycle restarts from point 1.

The QM/ELMO SCF procedure is also the starting point to carry out both post-HF/ELMO (e.g., MP2/ELMO or CCSD(T)/ELMO) ground state computations[79] and TDDFT/ELMO or EOM-CCSD/ELMO calculations for excited states[80]. The procedure has been implemented[78-80] by modifying the *Gaussain09* suite of programs.[103]

Finally, the IMOM/ELMO technique proposed in the present paper has been straightforwardly coded by changing the above-described QM/ELMO SCF algorithm through the substitution of the *aufbau* principle with the criterion based on the projectors (see equation (1)) for the occupation of the molecular orbitals of the QM subsystem.

III. COMPUTATIONAL DETAILS

In order to assess performances and capabilities of the new IMOM/ELMO method in dealing with excited states of large systems, we have initially performed test calculations on relatively large molecules (1-decanoic acid, 1-decene and (3E,5E,7E)-1,3,5,7,9-decapentaene (hereinafter simply indicated as decanoic acid, decene and decapentaene, respectively)) or group of molecules (acrolein surrounded by water).

Finally, to also prove usefulness and potentialities of the new approach, we applied it to a photobiology problem, namely the variation of the absorption spectrum of oxidized flavin mononucleotide (FMN) when it is bound to the protein flavodoxin.[104] The computational details for each of the above-mentioned test calculations are reported in the following subsections. For the sake of completeness, we already mention that all the computations were performed using the *Gaussian09* quantum chemistry package[103] in its standard version and also in our in-house variant.

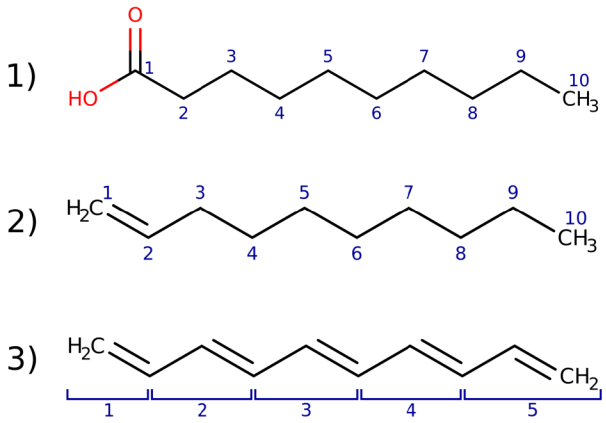


Figure 1. Schematic representation of the model systems taken into account for the initial validation of the IMOM/ELMO method: (1) decanoic acid, (2) decene, and (3) decapentaene. For (1) and (2), the reported numbers indicate the labels of the carbon atoms progressively included in the QM region for the IMOM/ELMO calculations, while, for (3), they represent the labels of the groups of atoms successively treated at quantum mechanical level.

III.A Test calculations on decanoic acid and decene. The geometries of decanoic acid, decene and decapentaene (see Figure 1) were initially optimized at B3LYP level with basis-set cc-pVDZ. The resulting geometries were afterwards used to compute fully IMOM and IMOM/ELMO calculations at B3LYP level with Pople basis-sets (6-31G(d), 6-311G(d), 6-31+G(d) and 6-311+G(d)) and with correlation-consistent sets of basis functions (cc-pVDZ, cc-pVTZ and aug-cc-pVDZ). Basis-set aug-cc-pVTZ was not taken into account due to convergence problems of the fully QM calculations

on the three test molecules, which prevented us to have benchmark values for the corresponding IMOM/ELMO computations.

For decanoic acid, we considered the two local transitions $n \rightarrow \pi^*$ and $\pi \rightarrow \pi^*$, while, for decene, we analyzed the local $\pi \rightarrow \pi^*$ excited state and the local double excitation $\pi^2 \rightarrow (\pi^*)^2$. Finally, to show the limitations of the proposed technique, we also investigated the delocalized $\pi \rightarrow \pi^*$ excited state of decapentaene. As explained in the Theory section, for both the fully IMOM and the IMOM/ELMO calculations, the guess/reference orbitals to determine the excited state single Slater determinants were selected by first inspecting the MOs obtained for the ground state and then promoting one or more electrons from an occupied orbital of type n (in case of an $n \rightarrow \pi^*$ excitation) or of type π (in case of the $\pi \rightarrow \pi^*$ and $\pi^2 \rightarrow (\pi^*)^2$ excitations) to a virtual orbital of π symmetry. In many cases, this corresponded to promoting one or more electrons from the HOMO or HOMO-1 to the LUMO of the considered systems. For the sake of completeness, the n, π and π^* selected orbitals for the IMOM/ELMO and the traditional IMOM calculations (with the 6-311+G(d) basis-set) on decanoic acid, decene and decapentaene are reported in the Supporting Information (see Figures S3-S5).

Concerning the IMOM/ELMO computations, the size of the QM region was gradually increased to evaluate the effect of the ELMO embedding on the description of the excited states. For decanoic acid, we considered from one to seven CH_2 alkyl moieties in the QM subsystem along with the carboxylic group; namely, we considered from two to eight carbon atoms (and relative substituents) in the QM region. For decene, we included from one to six alkyl moieties in the QM subsystem together with the ethylene group; in other words, we considered from three to eight carbon atoms (and relative substituents) in the QM region. For decapentaene, we took into account from

one to three propenyl-like units (i.e., from group 2 to group 4 in panel 3 of Figure 1) in the QM subsystem in addition to the first two carbon atoms and relative hydrogens (see group 1 in panel 3 of Figure 1). The results of the IMOM/ELMO calculations were compared to the corresponding fully IMOM ones. In particular, both excitation energies and oscillator strengths were analyzed.

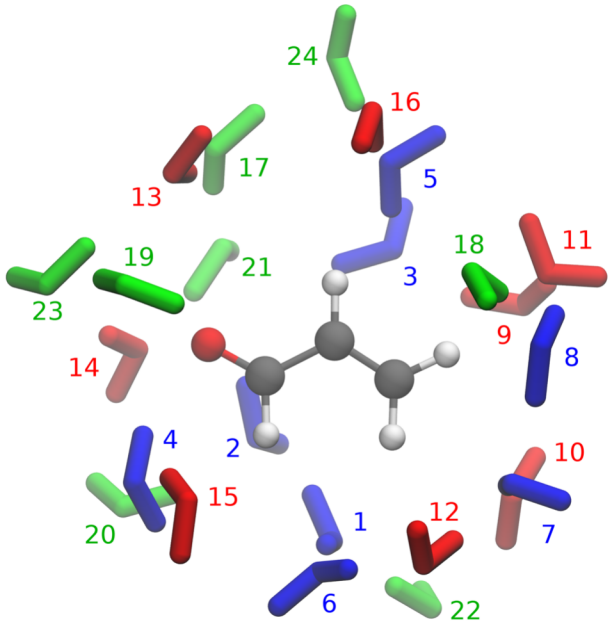


Figure 2. Solvated acrolein: each water molecule is labeled with a number that increases with the distance of the molecule from the acrolein barycenter. The eight closest water molecules (1-8) are depicted in blue, the second eight closest water molecules (9-16) are colored in red, while the eight farthest water molecules (17-24) are represented in green.

III.B Test calculations on solvated acrolein. The system chosen for this validation tests consisted of an acrolein molecule surrounded by 24 solvent water molecules (see Figure 2). All computations were performed at B3LYP level with basis-set aug-cc-pVDZ using a geometry taken from reference 75 and originally obtained through a Molecular Dynamics simulation. We considered excitations $n \rightarrow \pi^*$ and $\pi \rightarrow \pi^*$ and, for all the IMOM and IMOM/ELMO computations, we selected the guess/reference orbitals from the MOs resulting from the ground state calculations. We promoted one

electron from an n or π occupied orbital to a proper π^* virtual orbital to respectively obtain the single Slater determinants associated with the $n \rightarrow \pi^*$ and $\pi \rightarrow \pi^*$ excited states. Also in this case, the n, π and π^* orbitals selected for all the IMOM/ELMO and traditional IMOM calculations are reported in the Supporting Information (see Figures S6 and S7).

For the $n \rightarrow \pi^*$ excited state, a fully IMOM calculation was carried out on the complete system (i.e., acrolein plus 24 water molecules) and the obtained excitation energy and oscillator strength were used as reference values to check the convergence of the IMOM/ELMO computations performed by gradually enlarging the QM region, from a subsystem consisting only of acrolein and one (i.e., the closest) water molecule to a subsystem consisting of acrolein and 20 water molecules (see Figure 2). The water molecules were gradually introduced in the computations on the basis of their distance from the barycenter of acrolein.

Concerning the $\pi \rightarrow \pi^*$ excitation, fully IMOM calculations showed convergence problems when we considered a large amount of water molecules around acrolein. This is due to the fact that, as the number of water molecules increases, all orbitals of π symmetry significantly delocalize. It is thus impossible to select the correct MO from which to promote the electron and convergence problems arise. As already mentioned in the Introduction, this is one of the main motivations that prompted us to propose the IMOM/ELMO approach. Therefore, to check the capabilities of the IMOM/ELMO method also for the $\pi \rightarrow \pi^*$ excitation in the solvated acrolein system, the references for our IMOM/ELMO calculations were the excitation energy and the oscillator strength obtained through the fully IMOM computation with the largest possible number of solvent molecules with which we reached convergence (namely, 16 water molecules). Also in this case, we performed IMOM/ELMO calculations for

the $\pi \rightarrow \pi^*$ excited state by gradually increasing the QM region, from acrolein plus the closest water molecule to acrolein plus 12 water molecules (see Figure 2).

The previous validation tests allowed the determination of a consensus-number of solvent molecules (6 water molecules, see the Results and Discussion section) to be included in the QM region in order to achieve chemical accuracy (0.043 eV) for both the $n \rightarrow \pi^*$ and the $\pi \rightarrow \pi^*$ excitation energies with respect to the corresponding fully IMOM computations. Hence, to monitor the effects of the surrounding water molecules on the analyzed spectroscopic properties, we kept fixed the size of the QM region and we performed an additional series of IMOM/ELMO calculations to compute the $n \rightarrow \pi^*$ and the $\pi \rightarrow \pi^*$ excitation energies and the relative oscillator strengths by increasing the number of solvent molecules in the calculations (from 2 to 18, including 4 additional water molecules at each step) and by treating them at ELMO level.

III.C Application to oxidized flavin mononucleotide in flavodoxin. For the final photobiology test calculations, we considered a system consisting of the complex between oxidized FMN and the flavodoxin protein (see left panel of Figure 3) solvated by 7919 water molecules. Starting from a crystal structure (PDB code: 5NLL[105]), the system was initially relaxed through a Molecular Dynamics (MD) simulation (see Supporting Information for more details), from which 60 snapshots were randomly extracted. These configurations were used to perform the different types of IMOM/ELMO calculations that will be described below and that allowed the determination of conformationally-averaged absorption spectra corresponding to different cases.

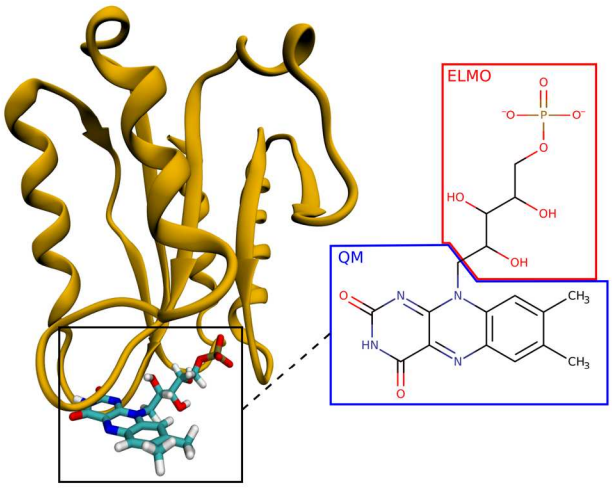


Figure 3. Complex between oxidized flavin mononucleotide (FMN) and the protein flavodoxin (left panel); zoom of the substrate with its subdivision into QM and ELMO regions. As indicated in the text, the ELMO subsystem also comprises protein atoms that are within 3.0 Å from any atom of FMN.

In particular, by exploiting the recent inclusion of the outer Molecular Mechanics layer in the framework of the QM/ELMO approach,[106] we initially carried out IMOM/ELMO/MM computations with (i) the QM region practically consisting only of the isoalloxazine triple ring of flavin mononucleotide (see right panel of Figure 3), (ii) the ELMO subsystem comprising the side chain of FMN (see again right panel of Figure 3) and all the atoms of flavodoxin within a 3.0 Å radius from any atom of the substrate, and (iii) the MM subunit corresponding to the rest of the protein and the surrounding water molecules. To evaluate the effects of long-range electrostatic interactions, the external MM region was then neglected, and simple IMOM/ELMO calculations were performed exploiting the same IMOM/ELMO partitioning scheme adopted for the above-described IMOM/ELMO/MM computations. Afterwards, the ELMO region was removed, and, always using geometries extracted from the MD production phase, we carried out *in vacuo* IMOM calculations only on the properly capped isoalloxazine triple ring. Furthermore, for the sake of completeness, and to better evaluate consequences and possible advantages of having an intermediate

buffer region treated at an approximate quantum mechanical level (i.e., the ELMO subsystem), we also performed simple IMOM/MM computations, with the isoalloxazine triple ring of FMN as the QM subunit and the rest of the system as the MM part.

Corresponding TDDFT/ELMO/MM, TDDFT/MM, TDDFT/ELMO and TDDFT calculations were also carried out on the same systems to obtain absorption spectra to be compared with the IMOM-based ones.

For all the above-described calculations, the QM region was always treated at PBE level, and the cc-pVDZ basis-set was adopted for the whole QM(/ELMO) region. When IMOM/ELMO/MM, TDDFT/ELMO/MM, IMOM/MM and TDDFT/MM calculations were performed, the MM subsystem was described through the ff14SB force field and the standard link atom strategy was employed to treat the ELMO/MM covalent boundaries.

III.D. ELMOs calculations. The ELMOs used in the QM/ELMO computations were previously obtained from the available ELMO libraries or from computations on proper model molecules carried out by exploiting a modified version of the GAMESS-UK software[107] that implements the Stoll equations[85,86] (see Supporting Information for more details about the ELMO theory). In particular, concerning decanoic acid and decene, the extremely localized molecular orbitals for the alkyl groups were determined on the butane molecule using a geometry optimized at B3LYP/cc-pVDZ level; pertaining to decapentaene, the ELMOs for the propenyl-like units were calculated on the B3LYP/cc-pVDZ optimized geometry of (3E,5E)-1,3,5-heptatriene; for the study on solvated acrolein, the only necessary ELMOs were those for the water molecules and were computed on a geometry always optimized at B3LYP/cc-pVDZ level; finally, for the investigation on the FMN-flavodoxin

complex, the extremely localized molecular orbitals for the protein were directly taken from the ELMO databanks[90] (cc-pVDZ basis-set), while those for the ELMO subsystem of FMN were determined on the optimized geometry (B3LYP/cc-pVDZ level) of a model molecule that suitably describes the chemical environment of the different fragments (see Figure S8 in the Supporting Information).

Before each QM/ELMO calculation, the pre-computed ELMOs were transferred using the *ELMOdb* program[90], namely the software that is associated with the ELMO databanks and that implements the strategy originally proposed by Philipp and Friesner for the rotation of strictly localized bond orbitals[88,108] (see Supporting Information for more details).

IV. RESULTS AND DISCUSSION

In this section the results of the above-described test calculations will be shown and analyzed to preliminarily evaluate capabilities and limitations of the IMOM/ELMO approach. In subsection IV.A we will discuss the results obtained for decanoic acid, decene and decapentaene, while, in subsection IV.B, we will focus on the case of acrolein surrounded by solvent water molecules, which can be considered as a more real and challenging situation. Finally, in subsection IV.C, we will discuss the application of the new IMOM/ELMO technique to investigate the optical properties of the complex between oxidized flavin mononucleotide and the protein flavodoxin.

IV.A Decanoic acid, decene and decapentaene. First of all, let us consider the results obtained for decanoic acid using the Pople basis-sets (see Figure 4). For the $n \rightarrow \pi^*$ excitation energies (see Figure 4A), already with only two carbon atoms (and relative substituents) in the QM subsystem, the IMOM/ELMO calculations provided values that are in excellent agreement with those obtained from the corresponding

fully IMOM computations, namely the discrepancies are always within chemical accuracy (0.043 eV). More precisely, when only two carbon atoms are included in the QM region, the calculations with basis-sets without diffuse functions provided results at the limit of chemical accuracy (especially in the 6-311G(d) case). On the contrary, the computations that used basis-sets with diffuse functions gave excitation energies that differ from the corresponding IMOM ones by less than 0.02 eV. By increasing the size of the QM region, the situation further improves, and the excitation energies rapidly converge to the benchmark IMOM values. In particular, already with three atoms in the QM subsystem, the discrepancies are lower than 0.01 eV for all the basis-sets. Concerning the oscillator strength for the $n \rightarrow \pi^*$ excitation (see Figure 4B), using the different sets of basis functions we obtained results that are quite stable and in very good agreement with the fully IMOM values regardless of the QM region size. The largest discrepancy (7×10^{-5} in absolute value) was observed for the IMOM/ELMO calculation with the 6-31G(d) basis-set and three carbon atoms in the quantum mechanical subsystem, corresponding to a relative error of 3.5%. However, also for the 6-31G(d) basis-set, the oscillator strengths converge to the fully IMOM benchmark results as more carbon atoms are treated at fully quantum mechanical level.

Pertaining to the $\pi \rightarrow \pi^*$ excitation energies (see Figure 4C), for all the basis-sets the discrepancies from the fully IMOM benchmark values generally decrease as the size of the QM subunit becomes larger. Furthermore, and more importantly, already with only two carbon atoms in the quantum mechanical region, the IMOM/ELMO calculations reproduce the corresponding fully IMOM values within chemical accuracy, with the largest deviation (0.012 eV in absolute value) observed for the 6-31G(d) basis-set. Similar trends were also observed for the values of the $\pi \rightarrow \pi^*$

oscillator strength (see Figure 4D), which remain quite stable as the dimension of the QM regions varies, but which anyway converge to the fully IMOM results as more and more carbon atoms are included in the fully quantum mechanical subsystem. As one should expect, the oscillator strength values for the *bright* $\pi \rightarrow \pi^*$ excitation are two order of magnitude greater than those for the *dark* $n \rightarrow \pi^*$ excitation.

Similar results and trends were obtained from the calculations carried out on decanoic acid with the correlation-consistent basis-sets, for both the $n \rightarrow \pi^*$ and the $\pi \rightarrow \pi^*$ excitations (see Figure S9 in the Supporting Information).

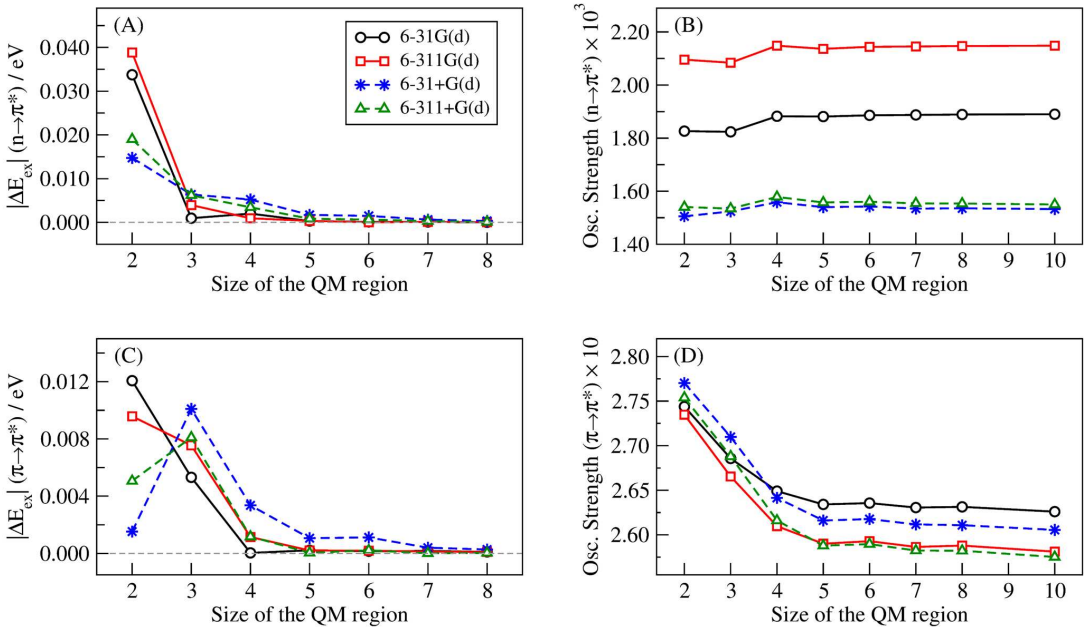


Figure 4. Results of fully IMOM and IMOM/ELMO calculations performed with Pople basis-sets on decanoic acid: (A) absolute discrepancies between the fully IMOM and IMOM/ELMO $n \rightarrow \pi^*$ excitation energies; (B) $n \rightarrow \pi^*$ oscillator strengths; (C) absolute discrepancies between the fully IMOM and IMOM/ELMO $\pi \rightarrow \pi^*$ excitation energies; (D) $\pi \rightarrow \pi^*$ oscillator strengths. For the oscillator strengths, the values obtained for 10 carbon atoms in the QM region are the fully IMOM references.

Now, let us consider the results obtained for decene with the Pople basis-sets (see Figure 5). Regarding the $\pi \rightarrow \pi^*$ excitation energy (Figure 5A), for all the considered sets of basis functions, the deviations from the corresponding fully IMOM values

decrease as the size of the QM region increases. Unlike decanoic acid, here for the smallest possible QM subsystem (i.e., three carbon atoms and relative substituents) the discrepancies are always above the chemical accuracy limit. However, already from the second smallest quantum mechanical region the deviations drop below the 0.043 eV threshold, with the largest absolute discrepancies observed when the 6-31G(d) and 6-31+G(d) basis-sets were used (0.036 eV). The tendencies for the $\pi \rightarrow \pi^*$ excitation energies are identical to those for the corresponding oscillator strengths (see Figure 5B), which converge to the fully IMOM results as we increase the number of carbon atoms in the quantum mechanical subunit.

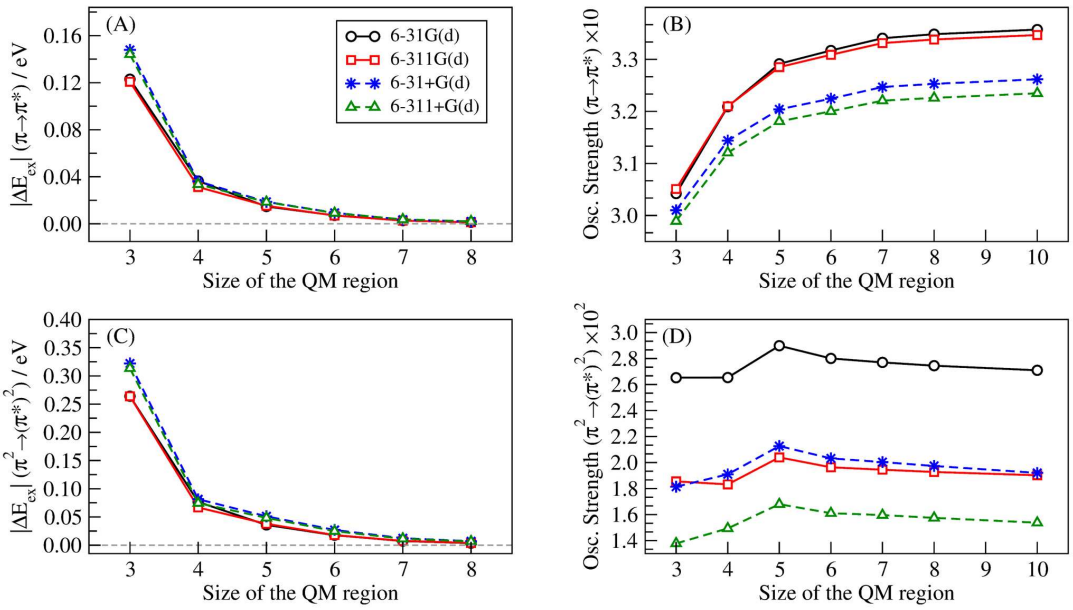


Figure 5. Results of fully IMOM and IMOM/ELMO calculations performed with Pople basis-sets on decene: (A) absolute discrepancies between the fully IMOM and IMOM/ELMO $\pi \rightarrow \pi^*$ excitation energies; (B) $\pi \rightarrow \pi^*$ oscillator strengths; (C) absolute discrepancies between the fully IMOM and IMOM/ELMO $\pi^2 \rightarrow (\pi^*)^2$ excitation energies; (D) $\pi^2 \rightarrow (\pi^*)^2$ oscillator strengths. For the oscillator strengths, the values obtained for 10 carbon atoms in the QM region are the fully IMOM references.

Also for the excitation energies associated with the $\pi^2 \rightarrow (\pi^*)^2$ excited state (see Figure 5C), we can notice that the IMOM/ELMO results clearly converge towards the

fully IMOM reference values as the size of the QM region increases. However, in these cases the convergence rates are slightly slower than those observed for the single excitation $\pi \rightarrow \pi^*$. In fact, for calculations performed with basis-sets without diffuse functions, the discrepancies start being lower than 0.043 eV when five carbon atoms (and relative substituents) are included in the quantum mechanical subsystem, while, for computations carried out by exploiting basis-sets without diffuse functions, chemical accuracy is achieved when six carbon atoms are treated quantum mechanically. Similar results can be observed also for the corresponding oscillator strengths (see Figure 5D), with convergence to the standard fully IMOM value that is practically always reached when at least five/six carbon atoms belong to the QM subunit.

As for decanoic acid, also for decene we obtained analogous results using the correlation-consistent basis-sets, for both the $\pi \rightarrow \pi^*$ and $\pi^2 \rightarrow (\pi^*)^2$ electronic transitions (see Figure S10 in the Supporting Information).

Finally, to show the limitations of the proposed method, we now report and discuss the IMOM/ELMO results obtained for the $\pi \rightarrow \pi^*$ excited state of decapentaene when the Pople basis-sets were adopted in the calculations (see Figure 6). Concerning the excitation energies (Figure 6A), we can see that, for all the sets of basis functions, the IMOM/ELMO values approach the fully IMOM benchmarks as the size of the QM region increases but, in these cases, without reaching the chemical accuracy threshold. For example, when the 6-31G(d) and 6-31+G(d) basis-sets were used, the IMOM/ELMO calculations with the largest quantum mechanical subsystem still respectively provided large discrepancies of 0.343 eV and 0.329 eV with respect to the corresponding fully IMOM values. This lack of convergence can be also noticed for the oscillator strengths (see Figure 6B) and in the computations with the

correlation-consistent sets of basis functions (see Figure S11 in the Supporting Information).

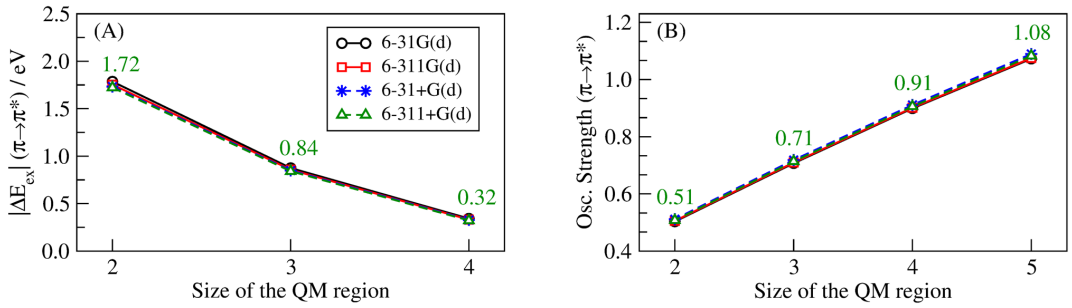


Figure 6. Results of fully IMOM and IMOM/ELMO calculations performed with Pople basis-sets on decapentaene: (A) absolute discrepancies between the fully IMOM and IMOM/ELMO $\pi \rightarrow \pi^*$ excitation energies; (B) $\pi \rightarrow \pi^*$ oscillator strengths. For the oscillator strengths, the values obtained with five groups of atoms in the QM region are the fully IMOM references (for the labels of the groups, see the lowest panel of Figure 1). The numerical values shown in the figure refer to the results of the calculations with the 6-311+G(d) basis-set; those associated with the other sets of basis functions are reported in Table S1 of the Supporting Information.

The observed results for decapentaene can be explained considering that the analyzed electronic transition is delocalized over the whole investigated system. This can be also evinced from the completely delocalized nature of the π^* molecular orbitals that resulted from standard QM ground state calculations and that were afterwards used in fully IMOM computations (see, for instance, the π^* MO obtained with the 6-311+G(d) basis-set and reported in the lowest panel of Figure S5 in the Supporting Information). On the contrary, the π^* molecular orbitals resulting from QM/ELMO calculations are less delocalized (see again Figure S5) and this led to the unsatisfactory results discussed above. This problem does not occur for the excitations in the other investigated systems (decanoic acid and decene), for which the delocalization extent of the QM/ELMO π^* MOs is always comparable to the one of

the π^* MOs obtained through fully QM computations (see Figure S3 and S4 in the Supporting Information).

Therefore, from the results of the test calculations discussed in the previous paragraphs, it clearly emerges that the IMOM/ELMO strategy perfectly works for local electronic transitions. On the contrary, it is a less suitable method for the treatment of delocalized excitations, although it is also worth pointing out that this feature is generally observed for any other embedding technique applied to the study of excited states, not only for the approach presented in this paper.

IV.B Solvated acrolein. As a more challenging testbed, we have afterwards decided to apply our method to the case of acrolein surrounded by water. First of all, we focused on the *dark* $n \rightarrow \pi^*$ excitation and we compared the results of our IMOM/ELMO computations with variable numbers of water molecules in the QM region to the outcome of the fully quantum mechanical IMOM calculation carried out on the complete systems consisting of acrolein and 24 water molecules. The obtained excitation energies and oscillator strengths are reported in Figure 7.

Concerning the excitation energies (see Figure 7A), we do not have a clear convergence trend and the discrepancies tend to oscillate as a function of the QM region size. However, regardless of the number of water molecules, the IMOM/ELMO calculations always provided results that agree within chemical accuracy with the fully IMOM one (corresponding to the case with 24 water molecules in Figure 7A). In particular, in most of the cases, the absolute differences from the benchmark fully IMOM excitation energy are between 1×10^{-3} and 7×10^{-3}eV (namely, between 0.02 and 0.16 kcal/mol), while the largest observed deviation is of 0.021 eV, which was obtained including four water molecules in the QM subsystem. Also for the oscillator strength (see Figure 7B) it is not possible to

find a clear trend. Nevertheless, according to the fact that the $n \rightarrow \pi^*$ excitation is dark, all the oscillator strength values are quite small (order of 10^{-5}) and, therefore, we believe that the observed oscillations are not really significant.

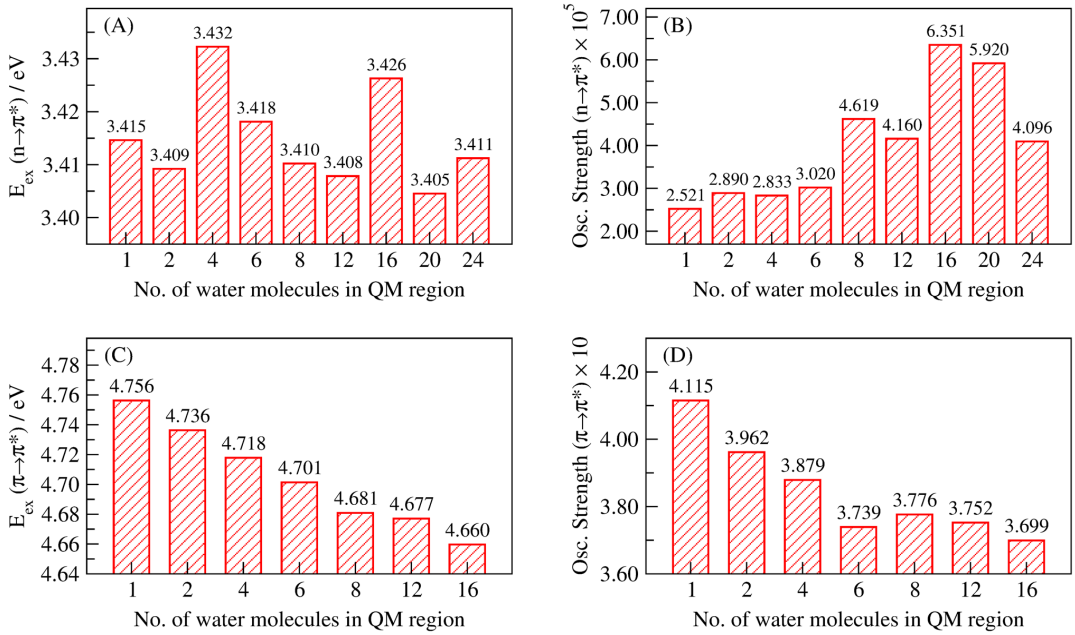


Figure 7. Results of IMOM/ELMO calculations (aug-cc-pVDZ basis-set) performed on solvated acrolein by increasing the size of the QM region: (A) $n \rightarrow \pi^*$ excitation energies; (B) $n \rightarrow \pi^*$ oscillator strengths; (C) $\pi \rightarrow \pi^*$ excitation energies; (D) $\pi \rightarrow \pi^*$ oscillator strengths. Computations with 24 and 16 water molecules in the QM region correspond to fully IMOM calculations for the $n \rightarrow \pi^*$ and $\pi \rightarrow \pi^*$ excitations, respectively.

We have afterwards considered the bright $\pi \rightarrow \pi^*$ excitation. As already explained in the Computational Details section, due to convergence problems we were not able to obtain the fully IMOM benchmark results for the complete system. For this reason, we here remind that, in this case, the benchmark excitation energy and oscillator strength refer to a fully IMOM calculation on a system constituted by acrolein and 16 water molecules (see again Figure 2). The results of the IMOM/ELMO and fully IMOM computations on this system for the $\pi \rightarrow \pi^*$ excitation are also reported in Figure 7.

Analyzing the excitation energies (see Figure 7C), the IMOM/ELMO values clearly converge to the fully IMOM one as we increase the number of water molecules in the quantum mechanical subsystem. Moreover, unlike what we observed for the $n \rightarrow \pi^*$ excitation, not all the discrepancies with respect to the IMOM values are within chemical accuracy and deviations drop below the 0.043 eV threshold from the IMOM/ELMO calculation with 6 water molecules in the QM region. Pertaining to the oscillator strengths (see Figure 7D), here it is possible to notice a clearer convergence compared to the $n \rightarrow \pi^*$ case. This is probably due to the fact that, since the $\pi \rightarrow \pi^*$ excitation is bright, the oscillator strength values are much larger and, therefore, they are less prone to subtle variations due to numerical fluctuations.

Finally, we decided to evaluate the effects of the solvent water molecules on the $n \rightarrow \pi^*$ and $\pi \rightarrow \pi^*$ excitations. As already mentioned in the section dedicated to Computational Details, to accomplish this task we performed a series of IMOM/ELMO calculations with only 6 water molecules in the QM region and a variable number of solvent molecules (from 2 to 18) in the ELMO subsystem. We chose to include 6 water molecules in the QM subunit because, from the above-discussed results, that amount of solvent molecules seems a good compromise to achieve chemical accuracy for both the $n \rightarrow \pi^*$ and the $\pi \rightarrow \pi^*$ excitations. The values of the obtained excitation energies and oscillator strengths are reported in Figure 8.

We can observe that the $n \rightarrow \pi^*$ excitation energy and oscillator strength have an almost analogous trend as a function of the number of water molecules considered in the computation and treated at ELMO level (see Figures 8A and 8B). In fact, the two quantities increase until 10 water molecules are included in the ELMO region and then decrease to values that are more similar to the ones observed for the

IMOM/ELMO calculation with only two water molecules in the ELMO subsystem. To check the correctness of these results, we performed additional TDDFT calculations (B3LYP/aug-cc-pVDZ and CAM-B3LYP/aug-cc-pVDZ levels) by gradually increasing the global number of water molecules (from 6 to 24) and we recovered the same qualitative trends described above (see Figure S12 in the Supporting Information). Concerning the $\pi \rightarrow \pi^*$ case, the excitation energy always increases as more water molecules are considered in the computations, although the values obtained with 10 and 14 water molecules in the ELMO region are practically identical (see Figure 8C). On the contrary, the oscillator strength slightly decreases going from 2 to 10 water molecules in the ELMO subsystem, but it increases when 14 and 18 solvent molecules are treated with frozen extremely localized molecular orbitals (see Figure 8D).

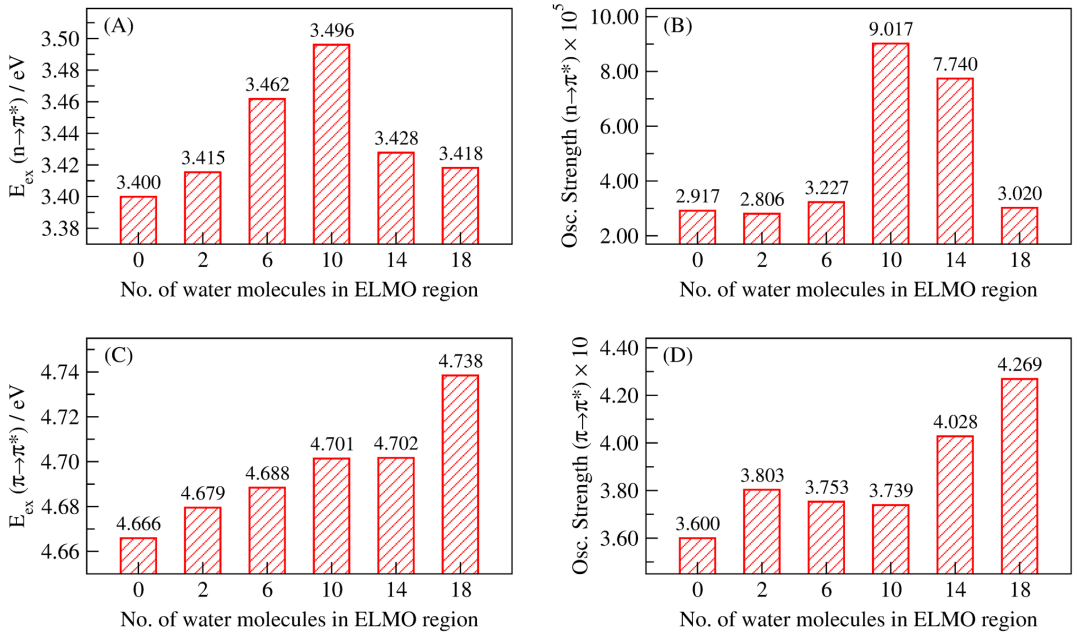


Figure 8. Results of IMOM/ELMO calculations (aug-cc-pVDZ basis-set) performed on solvated acrolein with six water molecules in the QM region and gradually increasing the number of water molecules in the ELMO region: (A) $n \rightarrow \pi^*$ excitation energies; (B) $n \rightarrow \pi^*$ oscillator strengths; (C) $\pi \rightarrow \pi^*$ excitation energies; (D) $\pi \rightarrow \pi^*$ oscillator strengths. Computations with zero water molecules in the ELMO region obviously correspond to fully IMOM calculations with only six solvent molecules surrounding acrolein.

IV.C FMN-flavodoxin complex. As final assessment of the novel IMOM/ELMO approach, we considered its application to the complex formed by oxidized flavin mononucleotide with the protein flavodoxin. We particularly evaluated how the absorption spectrum changes when the environment of the chromophore is progressively neglected in the calculations. The results of the IMOM-based computations are reported in Figure 9A. All the spectra are characterized by two peaks. However, when the long-range electrostatic interactions are turned off (IMOM/ELMO calculations instead of IMOM/ELMO/MM computations), we notice clear blue shifts for both the two electronic transitions, with the variation corresponding to the higher excitation energy that is clearly larger (0.09 eV against 0.04 eV). Further blue-shifts are observed when simple IMOM calculations are performed on the isolated saturated chromophore (namely, on the properly capped isoalloxazine triple ring). Also in this case the variation is significantly larger for the excited state of higher excitation energy. In fact, by comparing the IMOM/ELMO/MM results on the full systems to the IMOM ones on the capped chromophore, the ΔE_{ex} values amount to 0.07 eV and 0.28 eV for the first and the second peak, respectively. These results indicate that the environment has a non-negligible influence and, more importantly for this study, that its effects are captured by our new multi-scale IMOM-based approach with embedding provided by transferred and frozen extremely localized molecular orbitals.

Furthermore, as indicated in the section dedicated to the computational details, we also performed more standard IMOM/MM calculations. In the obtained absorption spectrum (see the turquoise curve in Figure 9A), we can observe that the two peaks are red-shifted by 0.06 and 0.07 eV compared to the IMOM/ELMO/MM case. This might be probably seen as a consequence of short-range QM/MM electrostatic

interactions due to presence of MM point charges that are very close to the QM region, a drawback that is avoided when the intermediate ELMO layer is used in the calculations.

To support the previous statements, we afterwards carried out additional TDDFT-based computations by exploiting the already tested TDDFT/ELMO embedding technique[80] and by following the same philosophy adopted for the above-discussed IMOM-based calculations. To construct all the TDDFT spectra (see Figure 9B), for each selected MD configuration we systematically considered all the excited states characterized by excitation energies lower than 3.5 eV and by non-negligible oscillator strengths.

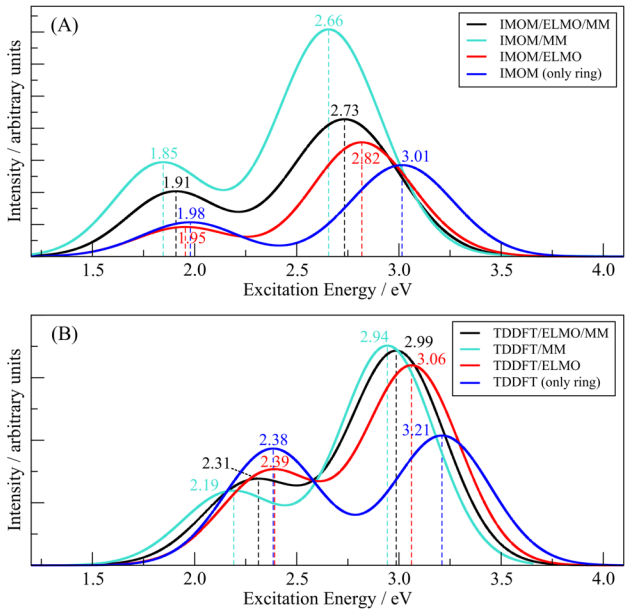


Figure 9. (A) IMOM-based and (B) TDDFT-based optical absorption spectra of oxidized flavin mononucleotide (FMN) in flavodoxin. The black curves represent the cases in which the full environment of the chromophore is fully taken into account at ELMO/MM level; the turquoise curves show the results of considering the full environment of the chromophore only at MM level; the red curves depict the situations in which only the closest part of the environment to the chromophore is taken into account and treated at ELMO level; the blue curves show the results of only considering the isoalloxazine triple ring in the calculations. All spectra were obtained by convolving Gaussian functions centered on different peaks with full-width-at-half-maximum (fwhm) of 0.5 eV.

By neglecting the long-range interactions, also in this case we observe blue shifts for the two absorption peaks. However, unlike what was seen for the IMOM calculations, the recorded ΔE_{ex} are approximately equivalent when the TDDFT/ELMO/MM results are compared to the TDDFT/ELMO ones (0.08 eV and 0.07 eV for the first and second electronic transition, respectively). Blue shifts are also noticed when even the closer ELMO region is not considered in the TDDFT computations, with variations of the excitation energies (compared to the TDDFT/ELMO/MM results) of 0.07 eV and 0.22 eV that, for this case, are completely in line with the corresponding IMOM-based results discussed above. Finally, we can also see that the standard TDDFT/MM calculations provided results analogous to the IMOM/MM ones, with the red-shifts for the two peaks in the spectrum amounting to 0.12 and 0.05 eV and again probably ascribable to the short-range electrostatic interactions between the QM and MM subsystems.

Therefore, the test calculations described in this subsection mainly showed that the proposed IMOM/ELMO(/MM) method is completely able to capture environment effects on absorption properties even in the case of biological systems, providing trends that are in perfect qualitative agreement with those resulting from the application of analogous and more standard TDDFT-based techniques.

V. CONCLUSIONS

In this work, we have presented the extension of the ΔSCF Initial Maximum Overlap Method (ΔSCF-IMOM) to the treatment of local excited states of large systems. In fact, although IMOM is much simpler and less computational expensive than multi-reference techniques for excited states, it cannot be always straightforwardly applied to systems of remarkable size. This is mainly due to the non-negligible delocalization

of the ground state MOs among which to select the guess/reference orbitals for the computation of the excited state wave function.

This goal has been achieved by coupling IMOM with the recently developed QM/ELMO technique, a fully quantum mechanical embedding approach that enables the treatment of only a small part of an investigated system at quantum mechanical level, while the rest is described through transferred and frozen extremely localized molecular orbitals. The new IMOM/ELMO strategy keeps the conceptual simplicity of the original Initial Maximum Overlap Method and enables easy investigations of localized excited states for large systems and macromolecules. In particular, from preliminary test calculations on relatively large systems, we observed that, for local electronic transitions, the new approach gives results that are in excellent agreement (i.e., within chemical accuracy) with the fully IMOM ones, but only treating a limited number of atoms at a fully quantum chemical level. Furthermore, the application of IMOM/ELMO to a photobiology problem showed that the proposed method is also able to provide results and trends that qualitatively agree with those resulting from more traditional computational techniques.

Further testbed calculations and improvements of the current IMOM/ELMO strategy are already envisaged. For example, exploiting the computational advantages offered by the QM/ELMO method when it is used in conjunction with post-HF techniques,[79] we are considering to introduce the description of electron correlation for excited states (e.g., at MP2 or Coupled Cluster level) within the IMOM/ELMO approach, as done for the original Maximum Overlap Method. Furthermore, another research direction offered by the new IMOM/ELMO technique could also consist in the possibility of computing X-ray absorption spectra (XAS) for large systems of biological interest. Moreover, to better asses the capabilities of the new method, we

are also planning to perform a future work where the new IMOM/ELMO technique and other strategies for the treatment of local excited states in large systems (e.g., all those based on our QM/ELMO philosophy and the projection-based embedding ones) will be compared in terms of chemical accuracy and computational cost.

Finally, it is worth noting that the strategy proposed in this work could be also easily combined with other kinds of ΔSCF algorithms (such as the more recent SGM and STEP approaches) or with the alternative ROKS method.

ASSOCIATED CONTENT

Supporting Information. Brief introduction on theory, transfer and libraries of extremely localized molecular orbitals (containing Figure S1 with examples of ELMOs and Figure S2 schematically showing reference frames and atomic triads required for the ELMOs rotation). Details about the preliminary orthogonalization procedure of the QM/ELMO method. Details on the Molecular Dynamics simulation of the solvated FMN-flavodoxin complex. Figures S3-S7 depicting n, π and π^* molecular orbitals selected for some of the IMOM/ELMO and fully IMOM calculations performed on decanoic acid, decene, decapentaene and solvated acrolein. Figure S8 showing the model molecule used for the computation of the extremely localized molecular orbitals for the side chain of FMN. Figure S9-S11 reporting the results of fully IMOM and IMOM/ELMO calculations performed with correlation-consistent basis-sets on decanoic acid, decene and decapentaene. Figure S12 showing the results of standard TDDFT computations (B3LYP/aug-cc-pVDZ and CAM-B3LYP/aug-cc-pVDZ levels) carried out on solvated acrolein with different amounts of surrounding water molecules. Tables S1 and S2 reporting the absolute discrepancies of the excitation energies and the oscillator strengths obtained through

IMOM/ELMO and fully IMOM calculations for the $\pi \rightarrow \pi^*$ excited state of decapentaene. The Supporting Information is available free of charge on the ACS Publications website.

AUTHOR INFORMATION

Notes

The authors declare no competing financial interests.

ACKNOWLEDGEMNTS

The French Research Agency (ANR) is gratefully acknowledged for financial support of this work through the Young Investigator Project *QuMacroRef* (Grant No. ANR-17-CE29-0005-01). The High-Performance Computing Center *EXPLOR* of the University of Lorraine is thanked for providing computing time through the projects 2019CPMXX0966, 2019CPMXX0886 and 2019CPMXX1332. Fabien Pascale (CNRS & University of Lorraine, France) is also acknowledged for the set-up and maintenance of our local cluster, which was used to perform most of the calculations reported in this paper. Thibaud Etienne (University of Montpellier, France) is thanked for providing us the program to produce the absorption spectra.

REFERENCES

[1] Helgaker, T.; Jørgensen, P.; Olsen, J. *Molecular Electronic-Structure Theory*; John Wiley and Sons: New York, 2000.

[2] Buenker, R. J.; Peyerimhoff, S. D. CI method for the study of general molecular potentials. *Theor. Chim. Acta* **1968**, *12*, 183-199.

[3] Roos, B. O.; Taylor, P. R.; Siegbahn, P. E. M. A complete active space SCF method (CASSCF) using a density matrix formulated super-CI approach. *Chem. Phys.* **1980**, *48*, 157-173.

[4] Roos, B. O. The complete active space SCF method in a fock-matrix-based super-CI formulation. *Int. J. Quantum Chem.* **1980**, *18*, 175-189.

[5] Andersson, K.; Malmqvist, P. A.; Roos, B. O.; Sadlej, A. J.; Wolinski, K. Second-order perturbation theory with a CASSCF reference function. *J. Phys. Chem.* **1990**, *94*, 5483-5488.

[6] Andersson, K.; Malmqvist, P. A.; Roos, B. O. Second-order perturbation theory with a complete active space self-consistent field reference function. *J. Chem. Phys.* **1992**, *96*, 1218−1226.

[7] Roos, B. O.; Andersson, K.; Fülscher, M. P.; Malmqvist, P. A.; Serrano-Andrés, L.; Pierloot, K.; Merchán, M. Multiconfigurational perturbation theory: applications in electronic spectroscopy. *Adv. Chem. Phys.* **2007**, *93*, 219-331.

[8] Hirao, K. Multireference Møller-Plesset method. *Chem. Phys. Lett.* **1992**, *190*, 374-380.

[9] Rowe, D. J. Equations-of-Motion Method and the Extended Shell Model. *Rev. Mod. Phys.* **1968**, *40*, 153-166.

[10] Mukherjee, D.; Mukherjee, P. K. A response-function approach to the direct calculation of the transition-energy in a multiple-cluster expansion formalism. Chem. Phys. **1979**, *39*, 325−335.

[11] Sekino, H.; Bartlett, R. J. A linear response, coupled-cluster theory for excitation energy. *Int. J. Quantum Chem.* **1984**, *26*, 255-565.

[12] Dalgaard, E.; Monkhorst, H. J. Some aspects of the time-dependent coupled-cluster approach to dynamic response functions. *Phys. Rev. A* **1983**, *28*, 1217-1222.

[13] Koch, H.; Jørgensen, P. Coupled cluster response functions. *J. Chem. Phys.* **1990**, *93*, 3333-3344.

[14] Del Bene, J. E.; Ditchfield, R.; Pople, J. A. Self-Consistent Molecular Orbital Methods. X. Molecular Orbital Studies of Excited States with Minimal and Extended Basis Sets. *J. Chem.Phys.* **1971**, *55*, 2236-2241.

[15] Foresman, J. B.; Head-Gordon, M.; Pople, J. A.; Frisch, M. J. Towards a systematic molecular orbital theory for excited states. *J. Phys. Chem.* **1992**, *96*, 135-149.

[16] Head-Gordon M.; Rico, R. J.; Oumi, M.; Lee, T. J. A doubles corrections to electronic excited states from configuration interaction in the space of single substitutions. *Chem. Phys. Lett.* **1994**, *219*, 21-29.

[17] Runge, E.; Gross, E. K. U. Density-Functional Theory for Time- Dependent Systems. *Phys. Rev. Lett.* **1984**, *52*, 997−1000.

[18] Petersilka, M.; Gossmann, U. J.; Gross, E. K. U. Excitation Energies from Time-Dependent Density-Functional Theory. *Phys. Rev. Lett.* **1996**, *76*, 1212−1215.

[19] Casida, M. E. In *Recent Advances in Density Functional Methods, Part I*; Dong, D. P., Ed.; Recent Advances in Computational Chemistry; World Scientific: Singapore, 1995; pp 155−192.

[20] Drew, A.; Head-Gordon, M. Single-reference ab initio methods for the calculation of excited states of large molecules. *Chem. Rev.* **2005**, *105*, 4009-4037.

[21] Tozer, D. J.; Handy, N. C. Improving virtual Kohn-Sham orbitals and eigenvalues: application to excitation energies and static polarizabilities. *J. Chem. Phys.* **1998**, *109*, 10180-10189.

[22] Casida, M. E.; Jamorski, C.; Casida, K. C.; Salahub, D. R. Molecular excitation energies to high-lying bound states from time- dependent density-functional response theory:

characterization and correction of the time-dependent local density approximation ionization threshold. *J. Chem. Phys.* **1998**, *108*, 4439-4449.

[23] Tozer, D. J.; Amos, R. D.; Handy, N. C.; Roos, B. O.; Serrano- Andreś, L. Does density functional theory contribute to the understanding of excited states of unsaturated organic compounds? *Mol. Phys.* **1999**, *97*, 859-868.

[24] Dreuw, A.; Head-Gordon, M. Failure of Time-Dependent Density Functional Theory for Long-Range Charge-Transfer Excited States: The Zincbacteriochlorin-Bacteriochlorin and Bacteriochlorophyll-Spheroidene Complexes. *J. Am. Chem. Soc.* **2004**, *126*, 4007-4016.

[25] Blase, X.; Duchemin, I.; Jacquemin, D.; Loos, P.-F. The Bethe-Salpeter Equation Formalism. *J. Phys. Chem. Lett.* **2020**, *11*, 7371-7382.

[26] Jones, R. O.; Gunnarsson, O. The density functional formalism, its applications and prospects. *Rev. Mod. Phys.* **1989**, *61*, 689-746.

[27] Hellman, A.; Razaznejad, B.; Lundqvist, B. I. Potential-energy surfaces for excited states in extended systems. *J. Chem. Phys.* **2004**, *120*, 4593-4602.

[28] Gavnholt, J.; Olsen, T.; Engelund, M.; Schiøtz, J. Δ self-consistent field method to obtain potential energy surfaces of excited molecules on surfaces. *Phys. Rev. B* **2008**, 075441.

[29] Gilbert, A. T. B.; Besley, N. A.; Gill, P. M. W. Self-Consistent Field Calculations of Excited States Using the Maximum Overlap Method (MOM). *J. Phys. Chem. A* **2008**, *112*, 13164-13171.

[30] Besley, N. A.; Gilbert, A. T. B.; Gill, P. M. W. Self-Consistent calculations of core excited states. *J. Chem. Phys.* **2009**, *130*, 124308.

[31] Barca, G. M. J.; Gilbert, A. T. B.; Gill, P. M. W. Hartree-Fock description of excited states of H_2. *J. Chem. Phys.* **2014**, *141*, 111104.

[32] Barca, G. M. J.; Gilbert, A. T. B.; Gill, P. M. W. Simple Models for Difficult Electronic Excitations. *J. Chem. Theory Comput.* **2018**, *14*, 1501-1509.

[33] Barca, G. M. J.; Gilbert, A. T. B.; Gill, P. M. W. Excitation Number: Characterizing Multiply Excited States. *J. Chem. Theory Comput.* **2018**, *14*, 9-13.

[34] Hait, D.; Head-Gordon, M. Orbital Optimized Density Functional Theory for Electronic Excited States. *J. Phys. Chem. Lett.* **2021**, *12*, 4517-1529.

[35] Levi. G.; Ivanov, A. V.; Jónsson, H. Variational calculations of excited states *via* direct optimization of the orbitals in DFT. *Faraday Discuss.* **2020**, *224*, 448.

[36] Levi. G.; Ivanov, A. V.; Jónsson, H. Variational Density Functional Calculations of Excited States via Direct Optimization. *J. Chem. Theory Comput.***2020**, *16*, 6968-6982.

[37] Hait, D.; Head-Gordon, M. Excited State Orbital Optimization via Minimizing the Square of the Gradient: General Approach and Application to Singly and Doubly Excited States via Density Functional Theory. *J. Chem. Theory Comput.* **2020**, *16*, 1699-1710.

[38] Carter-Fenk, K.; Herbert, J. M. State-Targeted Energy Projection: A Simple and Robust Approach to Orbital Relaxation of Non-Aufbau Self-Consistent Field Solutions. *J. Chem. Theory Comput.* **2020**, *16*, 5067-5082.

[39] Yamaguchi, K.; Jensen, F.; Dorigo, A.; Houk, K. A spin correction procedure for unrestricted Hartree-Fock and Møller-Plesset wavefunctions for single diradicals and polyradicals. *Chem. Phys. Lett.* **1988**, *149*, 537-542.

[40] Frank, I.; Hutter, J.; Marx, D.; Parrinello, M. Molecular dynamics in low-spin excited states. *J. Chem. Phys.* **1998**, *108*, 4060−4069.

[41] Filatov, M.; Shaik, S. A spin-restricted ensemble-referenced Kohn-Sham method and its application to diradicaloid situations. *Chem. Phys. Lett.* **1999**, *304*, 429−437.

[42] Kowalczyk, T.; Tsuchimochi, T.; Chen, P.-T.; Top, L.; Van Voorhis, T. Excitation energies and Stokes shifts from a restricted open- shell Kohn-Sham approach. *J. Chem. Phys.* **2013**, *138*, 164101.

[43] Jones, L. O.; Mosquera, M. A.; Schatz, G. C.; Ratner, M. A. Embedding Methods for Quantum Chemistry: Applications form Materials to Life Sciences. *J. Am. Chem. Soc.* **2020**,

142, 3281-3295.

[44] Warshel, A.; Levitt, M. Theoretical Studies of Enzymic Reactions: Dielectric, Electrostatic and Steric Stabilization of the Carbonium ion in the Reaction of Lysozyme. *J. Mol. Biol.* **1976**, *103*, 227-249.

[45] Field, M. J.; Bash, P. A.; Karplus, M. A Combined Quantum Mechanical and Molecular Mechanical Potential for Molecular Dynamics Simulations. *J. Comput Chem.* **1990**, *11*, 700-733.

[46] Gao, J. Methods and Applications of Combined Quantum Mechanical and Molecular Mechanical Potentials. In *Reviews in Computational Chemistry*; Lipkowitz, K. B.; Boyd, D. B., Eds.; VCH Publishers, Inc.: Weinheim, Germany, 1996; Vol. 7, pp 119-186.

[47] Gao, J. Hybrid Quantum and Molecular Mechanical Simulations: An Alternative Avenue to Solvent Effects in Organic Chemistry. *Acc. Chem. Res.* **1996**, *29*, 298-305.

[48] Senn, H. M.; Thiel, W. QM/MM Methods for Biomolecular Systems. *Angew. Chem., Int. Ed.* **2009**, *48*, 1198-1229.

[49] Knizia, G.; Chan, G. K.-L. Density matrix embedding: A strong coupling quantum embedding theory. *J. Chem. Theory Comput.* **2013**, *9*, 1428−1432.

[50] Bulik, I. W.; Chen, W.; Scuseria, G. E. Electron correlation in solids via density embedding theory. *J. Chem. Phys.* **2014**, *141*, 035140.

[51] Fornace, M. E.; Lee, J.; Miyamoto, K.; Manby, F. R.; Miller, T. F., III. Embedded mean-field theory. *J. Chem. Theory Comput.* **2015**, *11*, 568−580.

[52] Ye, H,-Z.; Van Voorhis, T. Atom-Based Bootstrap Embedding For Molecules. *J. Phys. Chem. Lett.* **2019**, *10*, 6368-6374.

[53] Cortona, P. Self-consistently determined properties of solids without band-structure calculations. *Phys. Rev. B* **1991**, *44*, 8454−8458.

[54] Wesolowski, T. A.; Shedge, S.; Zhou, X. Frozen-Density Embedding Strategy for Multilevel Simulations of Electronic Structure. *Chem. Rev.* **2015**, *115*, 5891-5928.

[55] Henderson, T. M. Embedding wave function theory in density functional theory. *J. Chem. Phys.* **2006**, *125*, 014105.

[56] Jacob, C. R.; Neugebauer, J.; Visscher, L. Software news and update: A flexible implementation of frozen-density embedding for use in multilevel simulations. *J. Comput. Chem.* **2008**, *29*, 1011−1018.

[57] Huang, C.; Pavone, M.; Carter, E. A. Quantum mechanical embedding theory based on a unique embedding potential. *J. Chem. Phys.* **2011**, *134*, 154110.

[58] Elliott, P.; Cohen, M. H.; Wasserman, A.; Burke, K. Density functional partition theory with fractional occupations. *J. Chem. Theory Comput.* **2009**, *5*, 827−833.

[59] Genova, A.; Ceresoli, D.; Pavanello, M. Periodic subsystem density-functional theory. *J. Chem. Phys.* **2014**, *141*, 174101.

[60] Mi, W.; Pavanello, M. Nonlocal Subsystem Density Functional Theory. *J. Phys. Chem. Lett.* **2020**, *11*, 272-279.

[61] Goodpaster, J. D.; Barnes, T. A.; Miller, T. F., III. Embedded density functional theory for covalently bonded and strongly interacting subsystems. *J. Chem. Phys.* **2011**, *134*, 164108.

[62] Manby, F. R.; Stella, M.; Goodpaster, J. D.; Miller, T. F., III. A simple, exact density-functional theory embedding scheme. *J. Chem. Theory Comput.* **2012**, *8*, 2564−2568.

[63] Lee, S. J. R.; Welborn, M.; Manby, F. R.; Miller, T. F., III. Projection-Based Wavefunction-in-DFT Embedding. *Acc. Chem. Res.* **2019**, *52*, 1359-1368.

[64] Culpitt, T.; Brorsen, K. R.; Hammes-Schiffer, S. Density functional theory embedding with the orthogonality constrained basis-set expansion procedure. *J. Chem. Phys.* **2017**, *146*, 211101.

[65] Chulhai, D. V.; Goodpaster, J. D. Improved Accuracy and Efficiency in Quantum Embedding though Absolute Localization. *J. Chem. Theory Comput.* **2017**, *13*, 1503-1508.

[66] Claudino, D.; Mayhall, N. J. Automatic Partition of Orbital Spaces Based on Singular Value Decomposition in the Context of Embedding Theories. *J. Chem. Theory Comput.* **2019**, *15*, 1053-1064.

[67] Sæther, S.; Kjærgaard, T.; Koch, H.; Høyvik, I.-M. Density-Based Multilevel Hartree-Fock Model. *J. Chem. Theory Comput.* **2017**, *13*, 5282-5290.

[68] Marrazzini, G.; Giovannini, T.; Scavino, M.; Egidi, F.; Cappelli, C.; Koch, H. Multilevel Density Functional Theory. *J. Chem. Theory Comput.* **2021**, *17*, 791-803.

[69] Chulhai, D. V.; Jensen, L. External orthogonality in subsystem time-dependent density functional theory. *Phys. Chem. Chem. Phys.* **2016**, *18*, 21032-21039.

[70] de Lima Batista, A. P.; de Oliveira-Filho, A. G. S.; Galembeck, S. E. Photophysical properties and the NO photorelease mechanism of a ruthenium nitrosyl model complex investigated using the CASSCF-in-DFT embedding approach. *Phys. Chem. Chem. Phys.* **2017**, *19*, 13860−13867.

[71] Koh, K. J.; Nguyen-Beck, T. S.; Parkhill, J. Accelerating Realtime TDDFT with Block-Orthogonalized Manby-Miller Embedding Theory. *J. Chem. Theory Comput.* **2017**, *13*, 4173-4178.

[72] Ding, F.; Tsuchiya, T.; Manby, F. R.; Miller, T. F., III. Linear-Response Time-Dependent Embedded Mean Field Theory. *J. Chem. Theory Comput.* **2017**, *13*, 4216-4227.

[73] Bennie, S. J.; Curchod, B. F. E.; Manby, F. R.; Glowacki, D. R. Pushing the Limits of EOM-CCSD with Projector-Based Embedding for Excitation Energies. *J. Phys. Chem. Lett.* **2017**, *8*, 5559−5565.

[74] Tölle, J.; Böckers, M.; Neugebauer, J. Exact subsystem time-dependent density-functional theory. *J. Chem. Phys.* **2019**, *150*, 181101.

[75] Wen, X.; Graham, D. S.; Chulhai, D. V.; Goodpaster, J. D. Absolutely Localized Projection-Based Embedding for Excited States. *J. Chem. Theory Comput.* **2020**, *16*, 385-398.

[76] Scholz, L.; Tölle, J; Neugebauer, J. Analysis of environment response effects on excitation energies within subsystem-based time-dependent density functional theory. *Int. J. Quantum Chem.* **2020**, *120*, e26213.

[77] Haldar, S.; Dutta, A. K. A Multilayer Approach to the Equation of Motion Coupled-Cluster Method for the Electron Affinity. *J. Phys. Chem. A* **2020**, *124*, 3947-3962.

[78] Macetti, G.; Genoni, A. Quantum Mechanics/Extremely Localized Molecular Orbital Method: A Fully Quantum Mechanical Embedding Approach for Macromolecules. *J. Phys. Chem. A* **2019**, *123*, 9420-9428.

[79] Macetti, G.; Wieduwilt, E. K.; Assfeld, X.; Genoni. A. Localized Molecular Orbital-Based Embedding Scheme for Correlated Methods. *J. Chem. Theory Comput.* **2020**, *16*, 3578-3596.

[80] Macetti, G.; Genoni, A. Quantum Mechanics/Extremely Localized Molecular Orbital Embedding Strategy for Excited States: Coupling to Time-Dependent Density Functional Theory and Equation-of-Motion Coupled Cluster. *J. Chem. Theory Comput.* **2020**, *16*, 7490-7506.

[81] Wieduwilt, E. K.; Macetti, G.; Genoni, A. Climbing Jacob's Ladder of Structural Refinement: Introduction of a Localized Molecular Orbital-Based Embedding for Accurate X-ray Determinations of Hydrogen Atom Positions. *J. Phys. Chem. Lett.* **2021**, *12*, 463-471.

[82] Macetti G.; Wieduwilt, E. K.; Genoni, A. QM/ELMO a Multi-Purpose Fully Quantum Mechanical Embedding Scheme Based on Extremely Localized Molecular Orbitals. *J. Phys. Chem. A* **2021**, *125*, 2709-2726.

[83] Assfeld, X.; Rivail, J.-L.; Quantum Chemical Computations on Parts of Large Molecules: the Ab Initio Local Self Consistent Field Method. *Chem. Phys. Lett.* **1996**, *263*, 100-106.

[84] Ferré, N.; Assfeld, A.; Rivail, J.-L. Specific Force Field Parameters Determination for the Hybrid Ab Initio QM/MM LSCF Method. *J. Comput. Chem.* **2002**, *23*, 610-624.

[85] Stoll, H.; Wagenblast, G.; Preuss, H. On the Use of Local Basis Sets for Localized Molecular Orbitals. *Theor. Chim. Acta* **1980**, *57*, 169−178.

[86] Fornili, A.; Sironi, M.; Raimondi, M. Determination of Extremely Localized Molecular Orbitals and Their Application to Quantum Mechanics/Molecular Mechanics Methods and to the Study of Intramolecular Hydrogen Bonding. *J. Mol. Struct. (THEOCHEM)* **2003**, *632*, 157-172.

[87] Sironi, M.; Genoni, A.; Civera, M.; Pieraccini, S.; Ghitti, M. Extremely Localized Molecular Orbitals: Theory and Applications. *Theor. Chem. Acc.* **2007**, *117*, 685-698.

[88] Meyer, B.; Guillot, B.; Ruiz-Lopez, M. F.; Genoni, A. Libraries of Extremely Localized Molecular Orbitals. 1. Model Molecules Approximation and Molecular Orbitals Transferability. *J. Chem. Theory. Comput.* **2016**, *12*, 1052-1067.

[89] Meyer, B.; Guillot, B.; Ruiz-Lopez, M. F.; Jelsch, C.; Genoni, A. Libraries of Extremely Localized Molecular Orbitals. 2. Comparison with the Pseudoatoms Transferability. *J. Chem. Theory. Comput.* **2016**, *12*, 1068-1081.

[90] Meyer, B.; Genoni, A. Libraries of Extremely Localized Molecular Orbitals. 3. Construction and Preliminary Assessment of the New Databanks. *J. Phys. Chem. A* **2018**, *122*, 8965-8981.

[91] Genoni, A; Sironi, M. A Novel Approach to Relax Extremely Localized Molecular Orbitals: the Extremely Localized Molecular Orbital-Valence Bond Method. *Theor. Chem. Acc.* **2004**, *112*, 254-262.

[92] Genoni, A.; Fornili, A.; Sironi, M. Optimal Virtual Orbitals to Relax Wave Functions Built Up with Transferred Extremely Localized Molecular Orbitals. *J. Comput. Chem.* **2005**, *26*, 827-835.

[93] Genoni, A.; Ghitti, M.; Pieraccini, S.; Sironi, M. A novel extremely localized molecular orbitals based technique for the one-electron density matrix computation. *Chem. Phys. Lett.* **2005**, *415*, 256-260.

[94] Genoni, A.; Merz. K. M., Jr.; Sironi, M. A Hylleras functional based perturbative technique to relax extremely localized molecular orbitals. *J. Chem. Phys.* **2008**, *129*, 054101.

[95] Sironi, M.; Ghitti, M.; Genoni, A.; Saladino, G.; Pieraccini, S. DENPOL: A new program to determine electron densities of polypeptides using extremely localized molecular orbitals. *J. Mol. Struct. (THEOCHEM)* **2009**, *898*, 8-16.

[96] Malaspina, L. A.; Wieduwilt, E. K.; Bergmann, J.; Kleemiss, F.; Meyer, B.; Ruiz-López, M.-F.; Pal, R.; Hupf, E.; Beckmann, J.; Piltz, R. O.; Edwards, A. J.; Grabowsky, S.; Genoni, A. Fast and Accurate Quantum Crystallography: from Small to Large, from Light to Heavy. *J. Phys. Chem. Lett.* **2019**, *10*, 6973-6982.

[97] Grabowsky, S.; Genoni, A.; Bürgi, H.-B. Quantum Crystallography. *Chem. Sci.* **2017**, *8*, 4159-4176.

[98] Genoni, A.; Bučinský, L.; Claiser, N.; Contreras-García, J.; Dittrich, B.; Dominiak, P. M.; Espinosa, E.; Gatti, C.; Giannozzi, P.; Gillet, J.-M.; Jayatilaka, D.; Macchi, P.; Madsen, A. Ø.; Massa, L. J.; Matta, C. F.; Merz, K. M., Jr.; Nakashima, P. N. H.; Ott, H.; Ryde. U.; Schwarz, K.; Sierka, M.; Grabowsky, S. Quantum Crystallography: Current Developments and Future Perspectives. *Chem. Eur. J.* **2018**, *24*, 10881-10905.

[99] Massa, L.; Matta, C. F. Quantum Crystallography: A perspective. *J. Comput. Chem.* **2018**, *39*, 1021-1028.

[100] Grabowsky, S.; Genoni, A.; Thomas, S. P.; Jayatilaka, D. The Advent of Quantum Crystallography: Form and Structure Factors from Quantum Mechanics for Advanced Structure Refinement and Wavefunction Fitting. In *21st Century Challenges in Chemical Crystallography 2 - Structural Correlations and Data Interpretation. Structure and Bonding*; Mingos, D. M. P., Rathby, P., Eds.; Springer: Berlin and Heidelberg, Germany, 2020; Vol. 186, pp 65-144.

[101] Genoni, A.; Macchi, P. Quantum Crystallography in the Last Decade: Developments and Outlooks. *Crystals* **2020**, *10*, 473.

[102] Macchi, P. The connubium between crystallography and quantum mechanics. *Crystallogr. Rev.* **2020**, *26*, 209-268.

[103] Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, Revision D.01; Gaussian, Inc., Wallingford, CT, USA, 2009.

[104] Cannuccia, E.; Pulci, O.; Del Sole, R.; Cascella, M. Optical properties of flavin mononucleotide: A QM/MM study of protein environment effects. *Chem. Phys.* **2011**, *389*, 35-38.

[105] Ludwig, M. L.; Pattridge, K. A.; Metzger, A. L.; Dixon, M. M.; Eren, M.; Feng, Y.; Swenson, R. P. Control of oxidation-reduction potential in flavodoxin from Clostridium beijerinckii: the role of conformation changes. *Biochemistry* **1997**, *36*, 1259-1280.

[106] Macetti, G.; Genoni, A. Three-Layer Multiscale Approach Based on Extremely Localized Molecular Orbitals to Investigate Enzyme Reactions. *J. Phys. Chem. A.*, under evaluation.

[107] Guest, M. F.; Bush, I. J.; van Dam, H. J. J.; Sherwood, P.; Thomas, J. M. H.; van Lenthe, J. H.; Havenith, R. W. A.; Kendrick, J. The GAMESS-UK Electronic Structure Package: Algorithms, Developments and Applications. *Mol. Phys.* **2005**, *103*, 719−747.

[108] Philipp, D. M.; Friesner, R. A. Mixed Ab Initio QM/MM Modeling Using Frozen Orbitals and Tests with Alanine Dipeptide and Tetrapeptide. *J. Comput. Chem.* **1999**, *20*, 1468-1494.