A QM/MM Approach Using the AMOEBA Polarizable Embedding: From Ground State Energies to Electronic Excitations

Daniele Loco, Etienne Polack, Stefano Caprasecca, Louis Lagardère, Filippo Lipparini, Jean-Philip Piquemal, Benedetta Mennucci

HAL Id: hal-02126716
https://hal.archives-ouvertes.fr/hal-02126716

Submitted on 12 May 2019

A QM/MM approach using the AMOEBA polarizable embedding: from ground state energies to electronic excitations

Daniele Loco,[†] Étienne Polack,[‡,¶] Stefano Caprasecca,[†] Louis Lagardère,[§,‡] Filippo Lipparini,[*,‖] Jean-Philip Piquemal,[*,‡,⊥] and Benedetta Mennucci[*,†]

[†]Dipartimento di Chimica e Chimica Industriale, Università di Pisa, via G. Moruzzi 13, I-56124 Pisa, Italy

[‡]Sorbonne Universités, UPMC Univ. Paris 06, UMR 7616, Laboratoire de Chimie Théorique, F-75005, Paris, France

[¶]Sorbonne Universités, UPMC Univ. Paris 06, UMR 7598, Laboratoire Jacques-Louis Lions, F-75005, Paris, France

[§]Sorbonne Universités, UPMC Univ. Paris 06, Institut du Calcul et de la Simulation, F-75005, Paris, France

[‖]Institut für Physikalische Chemie, Universität Mainz, Duesbergweg 10-14, D-55128 Mainz, Germany

[⊥]CNRS, UMR 7598 and 7616, F-75005, Paris, France

E-mail: flippari@uni-mainz.de; jpp@lct.jussieu.fr; benedetta.mennucci@unipi.it

Abstract

A fully polarizable implementation of the hybrid Quantum Mechanics/Molecular Mechanics approach is presented, where the classical environment is described through the AMOEBA polarizable force field. A variational formalism, offering a self-consistent

relaxation of both the MM induced dipoles and the QM electronic density is used for ground state energies and extended to electronic excitations in the framework of Time-Dependent Density Functional Theory combined with a state specific response of the classical part. An application to the calculation of the solvatochromism of the pyridinium N-phenolate betaine dye used to define the solvent ET30 scale is presented. The results show that the QM/AMOEBA model not only properly describes specific and bulk effects in the ground state but it also correctly responds to the large change in the solute electronic charge distribution upon excitation.

1   Introduction

The idea of studying an intrinsically quantum-mechanical (QM) process taking place in a complex environment by partitioning the whole system into a (smaller) subsystem (S) and the environment (E) has a long history in quantum chemistry. Within this framework, S is identified as the minimal subunit where the process occurs, and is described using an accurate albeit expensive level of theory, typically QM, whereas the environment is treated at a much lower detail, but sufficient to properly describe its effects on the process under study. Two alternative strategies are most commonly followed: one can either model the environment as a polarizable continuum[1,2] or employ an atomistic description of E, but introducing a Molecular Mechanics (MM) force field (FF).[3–7] The latter class of approaches is commonly referred to as QM/MM, and is widely used, particularly when electronic processes in complex environments are studied.

Since the first formulations, different QM/MM approaches have been proposed; in most of them the *effective* Hamiltonian defining the electronic properties of the QM subsystem in the presence of the MM system, is divided into a term describing the isolated S, and a term taking into account its electrostatic interaction with the classical environment through point charges centered on the MM atoms. This formulation, known as "electrostatic embedding" is nowadays the most common QM/MM formulation: it in fact allows one to include the

effects of the classical subsystem in the determination of the QM electronic density and all its related properties. What, however, this approach still misses is the possibility for E to polarize in response to the charge density (and its eventual changes) of the S subsystem. To include such mutual polarization effect, an extension of the QM/MM formulation beyond the electrostatic embedding is necessary and different alternative strategies have been proposed so far.[8–20] In this work, we will illustrate the theoretical development and the computational implementation of a novel polarizable QM/MM approach, based on the AMOEBA polarizable FF,[21–23] which we will refer to as QM/AMOEBA. Within this framework, the environment polarization is achieved through the use of atomic point dipoles, which are induced as a response to the electric field generated by the S system as well as the same MM sites bearing distributed multipoles up to quadrupoles. An approach to couple AMOEBA to a QM package has been recently proposed[24] in the framework of a not self-consistent procedure. In this contribution, for the first time in the context of AMOEBA, the global relaxation of the mutual QM/MM polarization is achieved by solving the QM/AMOEBA equations in a self-consistent manner, without approximations.

The implementation is further extended to describe electronically excited states. As a matter of fact, non polarizable QM/MM formulations can be easily extended to excited states as the "new" operator has a one-electron nature. When a polarizable embedding like AMOEBA is used instead, a difficulty appears when a linear response (LR) approach as that commonly used in the Time-Dependent Density Functional Theory (TD-DFT) is used. In the LR formulation, in fact, the whole spectrum of the excitations of interest is determined in a single step calculation by solving for the poles of the proper response function. When the same problem is formulated within a polarizable embedding, an additional contribution has to be taken into account, namely the dynamic response of the E subsystem. In the standard formulation of polarizable models (both in their continuum and MM formulations) such a response is calculated through the transition densities corresponding to the different excitations: the oscillating transition density of the S subsystem induces an oscillating polar-

ization in the E subsystem which creates an in-phase response acting back to the transition density.[25,26] This LR formulation has been shown to properly describe the dynamic environment effect in excitations involving bright states characterized by a large transition dipole moment.[27] The same formulation, however, lacks the capability of describing the relaxation of the environment in response to the changes in the QM density upon excitation: it is therefore not suitable to model excitations involving large changes such as Charge-Transfer (CT) like excitations. To overcome this shortcoming, various models have been introduced to recover a state-specific (SS) description of the response both within a continuum[28–30] or an MM formulation.[20] Here we adopt a perturbative correction which is exactly equivalent to the so-called corrected Linear Response (cLR) scheme originally developed for polarizable continuum models.[28] Within this framework, a relaxed density matrix is calculated for the excited state of interest and the corresponding excitation energy is corrected for the interaction with the proper induced dipole moments within the environment.

To show the potentialities of the method, we have selected a very well-known solvatochromic probe, namely the pyridinium N-phenolate betaine dye (or 'betaine-30'). Due to its large negative solvatochromism combined with a high solubility in many different solvents, it was proposed as a solvatochromic indicator for the determination of the solvent polarity, through an empirical scale called $E_{\mathrm{T}}(30)$.[31,32] From a computational point of view this molecule is really challenging for at least two reasons: (i) it is zwitterionic in its ground state and (ii) the excitation is related to an intramolecular charge transfer (CT) between the pyrimidine and phenolate moiety which acts to reduce the zwitterionic character. For such a kind of excitation, the solvation model not only needs to be able to properly describe both specific and bulk effects in the ground state but it also has to correctly respond to the large change in the electronic charge distribution from the ground to the excited state. We will show that both these two requirements are fully satisfied by our implementation of the QM/AMOEBA.

2 Theory and Implementation

2.1 The AMOEBA Force Field

One of the main characteristics of AMOEBA is the improved description of the electrostatics through atomic multipoles, up to the quadrupoles, placed on each classical atom. Polarization effects are included by using an induced point-dipole model within the smeared Thole's damping interaction scheme.[33] This is achieved by also providing the classical atoms with atomic polarizabilities. In general, the polarizability $\boldsymbol{\alpha}_i$ is a symmetric matrix; in practice, the scalar isotropic polarizabilities $\langle \alpha_i \rangle \equiv \frac{1}{3}\text{tr}(\boldsymbol{\alpha}_i)$ are used instead.

In a purely classical framework, at each polarizable site i, the static multipolar distribution generates an electric field $\vec{E}_i$ that induces an electric point-dipole moment, $\vec{\mu}_i$. The set $\{\vec{\mu}_i\}_{i=1}^{N_{\text{Pol}}}$ is the unknown of the polarization problem and the minimizer of the functional[34]

$$\mathcal{E}^{\text{Pol}} = -\sum_{i=1}^{N_{\text{Pol}}} E_i^\alpha \mu_i^\alpha + \frac{1}{2}\sum_{i=1}^{N_{\text{Pol}}} (\boldsymbol{\alpha}_i^{-1})^{\alpha\beta}\mu_i^\alpha\mu_i^\beta + \frac{1}{2}\sum_{i=1}^{N_{\text{Pol}}}\sum_{j\neq i}^{N_{\text{Pol}}} \mathcal{T}_{ij}^{\alpha\beta}\mu_i^\alpha\mu_j^\beta, \tag{1}$$

where indexes α and β generally indicate cartesian components, for which the Einstein's summation convention is assumed; E_i and μ_i are the electric field and the induced dipole, both at the atom site i. Note that we will generally assume that, while all the classical sites are characterized by multipole moments, not all of them will be polarizable, *i.e.*, associated with a polarizability. The number of classical and polarizable sites are indicated as N_{MM} and N_{Pol}, respectively, with $N_{\text{MM}} \geq N_{\text{Pol}}$. Further details can be found in previous works on the force field;[23,34] here, we only recall the expression for the (damped) dipole interaction tensor:

$$\mathcal{T}_{ij}^{\alpha\beta} = -\frac{\delta_{\alpha\beta}}{|\vec{r}_i - \vec{r}_j|^3}\lambda_3(u_{ij}) + 3\frac{|\vec{r}_i - \vec{r}_j|^\alpha|\vec{r}_i - \vec{r}_j|^\beta}{|\vec{r}_i - \vec{r}_j|^5}\lambda_5(u_{ij}). \tag{2}$$

Several damping schemes to avoid the well known *polarization catastrophe* can be found in the literature.[33,35–37] The scheme employed in AMOEBA makes use of an exponential

damping:

$$
\begin{cases}
\lambda_3(u_{ij}) & = 1 - e^{-a u_{ij}^3} \\[2mm]
\lambda_5(u_{ij}) & = 1 - (1 + e^{-a u_{ij}^3}) e^{-a u_{ij}^3} \quad .
\end{cases}
\tag{3}
$$

where $u_{ij} = r_{ij}/(\langle \alpha_i \rangle \langle \alpha_j \rangle)^{\frac{1}{6}}$ is the *effective distance* between two polarizable sites i and j.

In a more compact notation, $\boldsymbol{E} = \left(\vec{E}_1^T, \vec{E}_2^T \cdots \vec{E}_{N_{\mathrm{Pol}}}^T \right)^T$ and $\boldsymbol{\mu} = \left(\vec{\mu}_1^T, \vec{\mu}_2^T \cdots \vec{\mu}_{N_{\mathrm{Pol}}}^T \right)^T$ are the collections of the electric field and induced point dipoles at each polarizable site, while the symmetric, positive definite matrix $\boldsymbol{T}$, the size of which is $3N_{\mathrm{Pol}} \times 3N_{\mathrm{Pol}}$ and which is usually called polarization or interaction matrix, is defined as in Equation (4), where $(\mathcal{T})_{ij}$ is the 3×3 matrix defined as in Eq. (2).

$$
\boldsymbol{T} =
\begin{pmatrix}
\boldsymbol{\alpha}_1^{-1} & \mathcal{T}_{12} & \ldots & \mathcal{T}_{1N_{Pol}} \\
\mathcal{T}_{21} & \boldsymbol{\alpha}_2^{-1} & \ldots & \mathcal{T}_{2N_{Pol}} \\
\vdots & \vdots & \ddots & \vdots \\
\mathcal{T}_{N_{Pol}1} & \mathcal{T}_{N_{Pol}2} & \ldots & \boldsymbol{\alpha}_{N_{Pol}}^{-1}
\end{pmatrix}
\tag{4}
$$

Equation (1) can now be rewritten in matrix form as

$$
\mathcal{E}^{Pol} = \frac{1}{2} \boldsymbol{\mu}^{\dagger} \boldsymbol{T} \boldsymbol{\mu} - \boldsymbol{\mu}^{\dagger} \boldsymbol{E}
\tag{5}
$$

It is possible to find the polarization energy as the minimum of $\mathcal{E}^{Pol}$ with respect to the induced point dipoles,[34] which corresponds to the solution to the linear system

$$
\frac{\partial \mathcal{E}^{Pol}}{\partial \boldsymbol{\mu}} = \boldsymbol{T} \boldsymbol{\mu} - \boldsymbol{E} = 0.
\tag{6}
$$

This is the general formulation of the polarization problem, and is clearly variational. In the case of AMOEBA, however, the polarization energy is no longer a variational functional of the induced dipoles. This is due to the fact that two sets of induced dipoles are generated by two electric fields, differing for the scaling of the local interactions. Particularly, one set

6

of dipoles, indicated as $\boldsymbol{\mu}_{\mathrm{d}}$, is due to the so called "direct field" $\boldsymbol{E}_{\mathrm{d}}$, which is produced by multipoles placed on all the other classical sites. The other set, indicated as $\boldsymbol{\mu}_{\mathrm{p}}$, is induced by a "polarization field" $\boldsymbol{E}_{\mathrm{p}}$, where the contribution of the 1–2, 1–3 and 1–4 neighbors is scaled. A more detailed treatment on this point can be found in other works of some of us.[38] Here it suffices to say that the previous expression can be recast taking into account the different fields and induced dipoles. We recall that

$$\mathcal{E}_{\mathrm{A}}^{\mathrm{Pol}} = -\frac{1}{2}\boldsymbol{\mu}_{\mathrm{d}}^{\dagger}\boldsymbol{E}_{\mathrm{p}}, \tag{7}$$

is the proper expression for the AMOEBA polarization energy. Imposing the stationarity conditions for both sets of dipoles

$$\begin{cases} \boldsymbol{T}\boldsymbol{\mu}_{\mathrm{d}} = \boldsymbol{E}_{\mathrm{d}} \\ \boldsymbol{T}\boldsymbol{\mu}_{\mathrm{p}} = \boldsymbol{E}_{\mathrm{p}}, \end{cases} \tag{8}$$

where each set of induced dipoles is the variational minimizer of the corresponding energy functional, it is possible to reformulate the AMOEBA polarization energy as the combination of three variational expression, finally obtaining

$$\mathcal{E}_{\mathrm{A}}^{\mathrm{Pol}}(\boldsymbol{\mu}_{\mathrm{d}}, \boldsymbol{\mu}_{\mathrm{p}}) = \frac{1}{2}\boldsymbol{\mu}_{d}^{\dagger}\boldsymbol{T}\boldsymbol{\mu}_{p} - \frac{1}{2}\left(\boldsymbol{\mu}_{\mathrm{p}}^{\dagger}\boldsymbol{E}_{\mathrm{d}} + \boldsymbol{\mu}_{\mathrm{d}}^{\dagger}\boldsymbol{E}_{\mathrm{p}}\right) \tag{9}$$

2.2 An SCF-QM/AMOEBA formulation

The coupling of a classical description of the environment with SCF-based methods has already been discussed, especially for polarizable continuum solvation models in a variational scheme.[39,40] Generalizing those formulations to the QM/AMOEBA approach, the global variational energy functional can be written as the sum of three terms: (i) a purely QM energy functional, i.e., the SCF energy functional $\mathcal{E}^{\mathrm{QM}}$, (ii) a purely MM term given by the sum of the bonding, dispersion/repulsion, electrostatic and polarization terms according to

their definition within the AMOEBA FF; and (iii) a coupling term; namely we have

$$\mathcal{E}(\boldsymbol{P}, \boldsymbol{\mu}) = \mathcal{E}^{\mathrm{QM}}(\boldsymbol{P}) + \mathcal{E}^{\mathrm{MM}}(\boldsymbol{\mu}) + \mathcal{E}^{\mathrm{Coup}}(\boldsymbol{P}, \boldsymbol{\mu}) = \mathcal{E}^{\mathrm{QM}}(\boldsymbol{P}) + \mathcal{E}^{\mathrm{Env}}(\boldsymbol{P}, \boldsymbol{\mu}), \tag{10}$$

where in the RHS of the equation we added together the coupling and MM energy functionals in $\mathcal{E}^{\mathrm{Env}}$. Notice how the variational strategy, by introducing the global energy functional in Eq. (10), automatically takes into account the mutual polarization effects between the QM density and the induced dipoles.[39,41]

The variational environment term, $\mathcal{E}^{\mathrm{Env}}$, can be written as the sum of a constant MM contribution $\mathcal{E}^{\mathrm{FF}} + \mathcal{E}^{(0)}$ which does not depend on either the electronic density or the induced dipoles, a polarization energy, and a QM/MM coupling part. For the following manipulations, it is convenient to write separately the interaction of the QM density with the static multipoles as $\mathcal{E}^{\mathrm{QM/MM}}(\mathbf{P})$, and to add the interaction between the induced dipoles and the QM density to the polarization energy, from now on named $\widetilde{\mathcal{E}}^{\mathrm{Pol}}(\mathbf{P}, \boldsymbol{\mu})$; the resulting expressions of the various terms are:

$$\mathcal{E}^{\mathrm{Env}}(\boldsymbol{P}, \boldsymbol{\mu}) = \mathcal{E}^{\mathrm{FF}} + \mathcal{E}^{(0)} + \widetilde{\mathcal{E}}^{\mathrm{Pol}}(\mathbf{P}, \boldsymbol{\mu}) + \mathcal{E}^{\mathrm{QM/MM}}(\boldsymbol{P}) \tag{11}$$

$$\widetilde{\mathcal{E}}^{\mathrm{Pol}}(\boldsymbol{P}, \boldsymbol{\mu}) = \frac{1}{2}\boldsymbol{\mu}_d^\dagger \boldsymbol{T} \boldsymbol{\mu}_p - \frac{1}{2}\left(\boldsymbol{\mu}_{\mathrm{p}}^\dagger \boldsymbol{E}_{\mathrm{d}} + \boldsymbol{\mu}_{\mathrm{d}}^\dagger \boldsymbol{E}_{\mathrm{p}}\right) - \frac{1}{2}\left(\boldsymbol{\mu}_{\mathrm{p}} + \boldsymbol{\mu}_{\mathrm{d}}\right)^\dagger \mathbf{E}^{\mathrm{QM}}(\mathbf{P}) \tag{12}$$

$$\mathcal{E}^{\mathrm{QM/MM}}(\boldsymbol{P}) = \boldsymbol{q}^\dagger \boldsymbol{V}^{\mathrm{QM}}(\boldsymbol{P}) - \boldsymbol{\mu}_s^\dagger \mathbf{E}^{\mathrm{QM}}(\boldsymbol{P}) - \boldsymbol{\Theta}^\dagger \nabla \boldsymbol{E}^{\mathrm{QM}}(\boldsymbol{P}). \tag{13}$$

where q_i, $\vec{\mu}_{s,i}$ and $\boldsymbol{\Theta}_i$ are the fixed charges, dipoles and quadrupoles, respectively.

The electronic terms of $\mathcal{E}^{\mathrm{QM/MM}}$ in Eq. (13) can be expressed as functions of the one-

particle electron density matrix elements:

$$
\left\{
\begin{aligned}
V^{\mathrm{QM}}(\vec{r}_i; \boldsymbol{P}) &= \sum_{\mu\nu}^{N_{\mathrm{b}}} P_{\mu\nu} V_{\mu\nu}(\vec{r}_i) = -\sum_{\mu\nu}^{N_{\mathrm{b}}} P_{\mu\nu} \int_{\mathbb{R}^3} \frac{\chi_\mu(\vec{r})\chi_\nu(\vec{r})}{|\vec{r}-\vec{r}_i|}\, \mathrm{d}^3 r \\
\vec{E}^{\mathrm{QM}}(\vec{r}_i; \boldsymbol{P}) &= \sum_{\mu\nu}^{N_{\mathrm{b}}} P_{\mu\nu} E_{\mu\nu}(\vec{r}_i) = \sum_{\mu\nu}^{N_{\mathrm{b}}} P_{\mu\nu} \int_{\mathbb{R}^3} \frac{\chi_\mu(\vec{r})\chi_\nu(\vec{r})(\vec{r}-\vec{r}_i)}{|\vec{r}-\vec{r}_i|^3}\, \mathrm{d}^3 r \\
\vec{\nabla}\vec{E}^{\mathrm{QM}}(\vec{r}_i; \boldsymbol{P}) &= \sum_{\mu\nu}^{N_{\mathrm{b}}} P_{\mu\nu} G_{\mu\nu}(\vec{r}_i) = \sum_{\mu\nu}^{N_{\mathrm{b}}} P_{\mu\nu} \int_{\mathbb{R}^3} \chi_\mu(\vec{r})\chi_\nu(\vec{r})\mathcal{J}\vec{E}(\vec{r}_i)\, \mathrm{d}^3 r.
\end{aligned}
\right.
\tag{14}
$$

where the electron density $\boldsymbol{P}$ has been expanded in a basis of atomic orbitals and $\mathcal{J}$ is the Jacobian of the electric field with elements $(\mathcal{J}\vec{E})_{\alpha\beta}(\vec{r}) = \frac{\partial E^\alpha(\vec{r})}{\partial r_\beta}$, being α and β generic Cartesian components. The corresponding terms induced by the nuclear distribution are trivial and are not reported.

Imposing the stationarity conditions for the global functional in Eq. ((10)), taking into account the constraints on the electronic density matrices, gives the coupled QM/AMOEBA equations.

The QM/AMOEBA Fock matrix $\widetilde{\boldsymbol{F}}$ is obtained by differentiating Eq. 10 with respect to the density matrix:

$$
\widetilde{\boldsymbol{F}} = \frac{\partial \mathcal{E}(\boldsymbol{P}, L, \boldsymbol{\mu})}{\partial \boldsymbol{P}} = \frac{\partial \mathcal{E}^0(\boldsymbol{P})}{\partial \boldsymbol{P}} + \frac{\partial \mathcal{E}^{\mathrm{Env}}(\boldsymbol{P}, \mu)}{\partial \boldsymbol{P}} = \boldsymbol{F}^{(0)} + \boldsymbol{F}^{\mathrm{Env}}.
\tag{15}
$$

where the term $\boldsymbol{F}^{(0)}$ of (15) corresponds to the Fock matrix of the pure QM electronic problem, while $\boldsymbol{F}^{\mathrm{Env}}$ is the contribution to $\widetilde{\boldsymbol{F}}$ from the classical environment. The elements of the latter can be written as

$$
F_{\mu\nu}^{\mathrm{Env}} = \boldsymbol{q}^\dagger \boldsymbol{V}_{\mu\nu} - \boldsymbol{\mu}_s^\dagger \boldsymbol{E}_{\mu\nu} - \boldsymbol{\Theta}^\dagger \boldsymbol{G}_{\mu\nu} - \frac{1}{2}(\boldsymbol{\mu}_{\mathrm{p}} + \boldsymbol{\mu}_{\mathrm{d}})^\dagger \boldsymbol{E}_{\mu\nu}
\tag{16}
$$

where the matrices $\boldsymbol{V}_{\mu\nu}$, $\boldsymbol{E}_{\mu\nu}$ and $\boldsymbol{G}_{\mu\nu}$ are those appearing in Eq. (14). Note that the $E_{\mu\nu}$ elements involved in the second term on the RHS of Eq. (16) are formally identical to those appearing in the first one, but are computed over a different set of atomic sites (the

9

polarizable ones only).

The linear equations for the dipoles are obtained by setting the derivatives of Eq. (10) with respect to both sets of dipoles to zero:

$$\boldsymbol{T}\boldsymbol{\mu}_{\mathrm{p}} = \boldsymbol{E}_{\mathrm{p}} + \boldsymbol{E}^{\mathrm{QM}}(\boldsymbol{P})$$

$$\boldsymbol{T}\boldsymbol{\mu}_{\mathrm{d}} = \boldsymbol{E}_{\mathrm{d}} + \boldsymbol{E}^{\mathrm{QM}}(\boldsymbol{P}). \tag{17}$$

Notice that the Fock matrix depends on the induced dipoles and that the RHSs of the dipole equations depend on the density matrix: the two sets of equations are therefore coupled, accounting correctly for the mutual polarization of the QM and classical charge densities. The coupled equations need to be solved iteratively, which is not a problem in practice as the SCF equations are already solved with an iterative algorithm. We note that the matrix $\boldsymbol{T}$ depends only on geometrical parameters, and can be computed and inverted at the first SCF cycle, thus reducing the computational requirements at the following steps. Alternatively, this problem can be also efficiently solved with an iterative procedure (see Section 2.4 for more details).

2.3 Extension to electronic excitations

The discussion so far has involved the energies and properties of the electronic ground state. To extend the treatment to excited states, we follow a linear response theory for SCF methods. A complete derivation of the polarizable QM/MM LR response equations can be found elsewhere,[14,41] here it suffices to say that the electronic transition energies ω_K and densities $\mathbf{X}_K, \mathbf{Y}_K$ are found by solving the modified Casida's equations[42]

$$\begin{pmatrix} \widetilde{\boldsymbol{A}} & \widetilde{\boldsymbol{B}} \\ \widetilde{\boldsymbol{B}} & \widetilde{\boldsymbol{A}} \end{pmatrix} \begin{pmatrix} \boldsymbol{X}_K \\ \boldsymbol{Y}_K \end{pmatrix} = \omega_K \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & -\mathbf{1} \end{pmatrix} \begin{pmatrix} \boldsymbol{X}_K \\ \boldsymbol{Y}_K \end{pmatrix}. \tag{18}$$

10

The matrices $\widetilde{A}$ and $\widetilde{B}$ are defined as

$$\widetilde{A}_{ai,bj} = \delta_{ab}\delta_{ij}(\epsilon_a - \epsilon_i) + (ai|bj) - c_x(ab|ij) + c_l f^{xc}_{ai,bj} + C^{\text{Pol}}_{ai,bj} \tag{19}$$

$$\widetilde{B}_{ai,bj} = (ai|bj) - c_x(aj|ib) + C^{\text{Pol}}_{ai,bj},$$

where ϵ are the orbital energies, $(pq|rs)$ are two-electron integrals in Mulliken notation and we assume that the orbitals are real and the coefficients c_x and c_l define whether we are considering the Hartree–Fock theory ($c_x = 1, c_l = 0$), pure DFT ($c_x = 0, c_l = 1$) or hybrid DFT. The elements of matrix $\boldsymbol{C}^{\text{Pol}}$ are due to the environment polarization:

$$C^{\text{Pol}}_{ai,bj} = -\sum_P^{N_{\text{Pol}}} \left(\int_{\mathbb{R}^3} \phi_a(\vec{r}) \frac{(\vec{r} - \vec{r}_p)}{|\vec{r} - \vec{r}_p|^3} \phi_i(\vec{r}) \mathrm{d}\vec{r} \right) \cdot \vec{\mu}^T_P(\phi_b, \phi_j). \tag{20}$$

The matrix C^{Pol} depends on the dipoles $\boldsymbol{\mu}^T$ induced by the transition density $\mathbf{P}^T_K = \mathbf{X}_K + \mathbf{Y}_K$, which are obtained by solving, for each couple of transition vectors $\mathbf{X}_K, \mathbf{Y}_K$, the response linear equations

$$\mathbf{T}\boldsymbol{\mu}^T = \mathbf{E}(\mathbf{P}^T_K), \tag{21}$$

where

$$\vec{E}_P(\mathbf{P}^T_K) = \sum_{ai} \mathbf{P}^T_{K,ai} \int_{\mathbb{R}^3} \phi_a(\vec{r}) \frac{(\vec{r} - \vec{r}_p)}{|\vec{r} - \vec{r}_p|^3} \phi_i(\vec{r}) \mathrm{d}\vec{r}. \tag{22}$$

Notice that in the right hand side of eq. 21 only the electric field due to the transition density appears: as the classical multipoles do not contribute to the transition dipoles, the p and d response dipoles introduced in section 2.1 are identical. Again, the modified Casida's equations depend (linearly) on the transition induced dipoles, which in turn depend (linearly) on the transition vectors $\mathbf{X}, \mathbf{Y}$: although the coupling is linear, it is still convenient to solve the couple equations iteratively, by solving the transition dipoles equations at each step of the iterative procedure used to solve Casida's equations.

The great advantage of the LR scheme is that it allows one to obtain a whole spectrum of transition energies at once. If the system is isolated, it provides the same results than state

specific (SS) approaches, where instead the wavefunction of each excited state is explicitly computed. By contrast, the introduction in the Hamiltonian of a non-linear term due to the presence of a polarizable environment results in a formal difference between the two approaches.[25,43] For this reason, a correction to the LR scheme (called *corrected* Linear Response, c-LR) has been proposed.[28] The difference between the two approaches (LR and c-LR) can be depicted by a two-step process for excitation in solution: first the molecule in its ground state, in equilibrium with the solvent, is excited to the state K in the presence of a solvent response frozen to the one of the solute ground state. The description for this step of the process is the same in both the approaches. In the second step, the fast electronic degrees of freedom of the solvent equilibrate with the electron density of the solute excited state, and in this part of the process the LR framework lacks in taking explicitly into account the energy variation accompanying this relaxation, whereas it accounts for a correction which, being originated by the dynamic solute- solvent interactions, can be classified as a part of dispersion. The working equation, namely

$$\omega_K^{c-LR} = \omega_K^0 - \frac{1}{2} \sum_P^{N_{pol}} \vec{\mu}_P(\vec{r}_P; \mathbf{P}_K^{\Delta}) \vec{E}(\vec{r}_P, \mathbf{P}_K^{\Delta}) \tag{23}$$

corresponds to the expression of the c-LR transition energy for the K-th excited state. In this formulation ω_K^0 represents the response of the system in a solvent frozen in its initial configuration (corresponding to the system in its ground-state), and it is the solution of a non-Hermitian eigensystem as in Eq. (18), where the $C_{ai,bj}^{\mathrm{Pol}} = 0$ but the orbitals and their corresponding energies, employed to build the A and B matrices have been obtained by solving the SCF equation for the solvated system. The $\mathbf{P}_K^{\Delta}$ is the so called *relaxed-density matrix*, computed through the so-called $\mathbf{Z}$-vector approach[44] as

$$\mathbf{P}_K^{\Delta} = \mathbf{P}_K^T + \mathbf{Z}_K \tag{24}$$

where $\mathbf{P}_K^T$ is the unrelaxed density matrix with elements given in terms of the transition

12

vectors $|\boldsymbol{X}_K, \boldsymbol{Y}_K\rangle$, and the $\boldsymbol{Z}$-vector contribution $\boldsymbol{Z}_K$ accounts for orbital relaxation effects.

2.4 Implementation

The hybrid QM/AMOEBA method has been implemented in a development version of the Gaussian software.[45] In the present implementation, the program computes all the terms that depend on the QM density, including the sets of induced dipoles. The $\mathcal{E}^{FF} + \mathcal{E}^{(0)}$ energy terms depend neither on the QM density nor on the induced dipoles and, in a static picture, are constant.

In order to compute the induced dipoles, both an inversion approach and an iterative one have been implemented. In the former case the polarization matrix (Eq. 4) is computed and inverted only once during the first step of the SCF, then stored and multiplied for the appropriate field to obtain the solutions to the polarization problem. At each SCF cycle only two matrix-vector products (one per set of dipoles) need to be performed, concentrating the computational effort almost entirely in the first SCF step.

By contrast, in the iterative approach, the polarization equations are solved iteratively at each SCF step or at each step of the iterative solution of Casida's equations, using as a guess the induced dipoles of a previous step, when available. Two algorithms have been implemented: Jacobi Iterations coupled with direct inversion in the iterative subspace (JI/DIIS) and the Preconditioned Conjugate Gradient (PCG).[34,46] Both are available for the ground and LR excited state problems.

The advantage of the inversion approach is that the most expensive part of the procedure, *i.e.*, the matrix inversion, needs only be performed once. However, the computational cost of such an operation grows with the cube of the number of polarizable sites: for large systems, iterative alternatives become mandatory. Furthermore, iterative methods can be easily combined with fast summation techniques, such as the Fast Multipole Method,[46,47] in order to achieve linear scaling in computational cost for the solution of the polarization equations. We here present the relative time requirements for the computation of the induced dipoles

in the whole SCF procedure, using the inversion and iterative algorithms. The calculations have been performed on the test system whose results will be discussed in Section 3. The dimension of the polarization problem has been artificially varied, by considering various polarization cutoffs, *i.e.*, by including a different number of induced dipoles. The dimension of the QM system, as well as that of the non-polarizable classical environment, have been kept fixed. Figure 1 reports the dependence of the relative times on the number of induced dipoles.
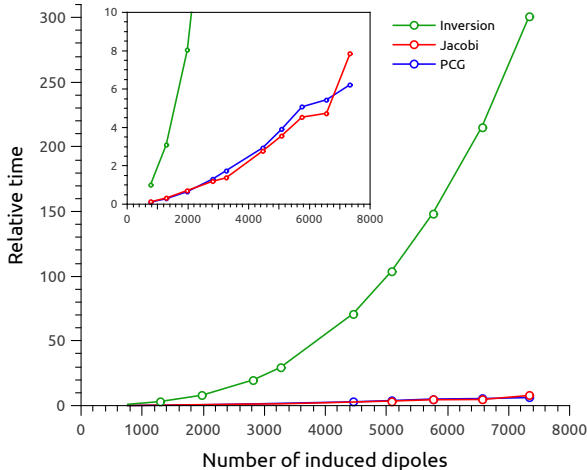
Figure 1: Relative times needed to solve for the induced dipoles, using different methods. Green line: inversion; Red line: Jacobi Iterations; Blue line: Preconditioned Conjugate Gradient. The times are cumulative and refer to the sum of all SCF cycles until convergence is reached. The number of SCF cycles considered is the average one across the various structures considered. The values have been normalized for comparison with respect to the time required to solve the smallest polarisation problem (7 Å polarization radius, 765 induced dipoles) with the inversion procedure.

From the plots reported in the figure, it is clear that the iterative procedures are more convenient than the inversion one. This becomes more evident as the size of the problem increases, since both JI and PCG algorithms scale quadratically with the number of induced dipoles, while the inversion algorithm, LU decomposition, is characterized by a cubic scaling. In particular, when the polarization radius is increased from 7 to 20 Å (number of dipoles increasing by $\sim 10\times$ from 765 to 7338), the inversion algorithm increases by $\sim 300\times$. For both JI and PCG algorithms, the relative increase is $\sim 60\times$.

Furthermore when the number of polarizable dipoles is small (polarization radius 7 Å, 765 dipoles), even if the JI/PCG methods are faster than the inversion one, the choice of the solution scheme does not really matter since the relative time to solve for the induced dipoles with respect to the total time is, for each of the three cases, less than the 0.1%. This is no longer true in the most expensive case (polarization radius 20 Å, 7338 dipoles), where almost $16, 5\%$ of the time for the calculation is spent in computing the induced dipoles with the inversion method (the iterative solutions take less than 1% of the total time).

3   A test case: the solvatochromism of betaine-30

In this section we present an application of the QM/AMOEBA implementation to the simulation of the excitation energies of a well-known solvatochromic probe, the pyridinium N-phenolate betaine dye, from now on indicated as "betaine-30" (see Figure 2). The solvent we selected for this test case is water, since it constitutes one of the most interesting cases due to its high polarity combined with an hydrogen-bonding character. In particular, we will try to dissect these two components of the solvent effect by comparing QM/AMOEBA with a continuum description using the Polarizable Continuum Model (PCM) within its Integral Equation Formalism (IEF).[48] We recall that this model describes the environment as a structureless continuum, characterized by its macroscopic dielectric function. A cavity containing the solute is built around it, and the solvent polarization as a response to the solute charge density is represented by an induced surface charge distribution on the cavity.

In order to obtain a realistic sampling of the solute–solvent interaction, a well-established procedure has been applied, which makes use of snapshots extracted from a classical molecular dynamics (MD) simulation with the AMBER General Force Field for organic molecules. An optimized geometry of the dye was solvated with a cubic box containing 11380 water molecules, described at TIP3P level,[49] with dimensions $74 \times 72 \times 78$ Å. The system was heated from 0 to 300 K for 100 ps with a 2 fs time step, employing the SHAKE algorithm[50]
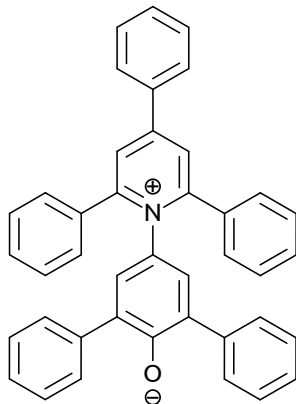
Figure 2: The betaine dye studied.

as in all the following steps. The Berendsen thermostat[51] was used to control the temper-
atue. A 200 ps equilibration in the NVT ensemble was then performed. Before the actual
production, the system was further equilibrated in the NPT ensamble at 1 atm for 5 ns.
Here, the Monte Carlo barostat implemented in Amber14[52] was employed. In the 30-ns
long MD simulation, the betaine-30 was kept frozen in its ground state equilibrium geome-
try (computed at QM/PCM level). A set of 100 uncorrelated snapshots was extracted, on
which the QM/AMOEBA calculations were performed. Since the effect of polarization is
short-ranged, and the computational cost increases markedly with the number of induced
dipoles, only the classical atoms within a certain radius (the *polarization radius*, R_{pol}) were
allowed to polarize. The optimal value of the radius was chosen after performing a conver-
gence test, where the first three excited states of the solvated betaine dye were computed at
increasing values of R_{pol}. The results, reported in Figure 3, show that convergence can be
observed starting from 10–12 Å. A safe value $R_{\text{pol}} = 15$ Å was chosen and employed in all
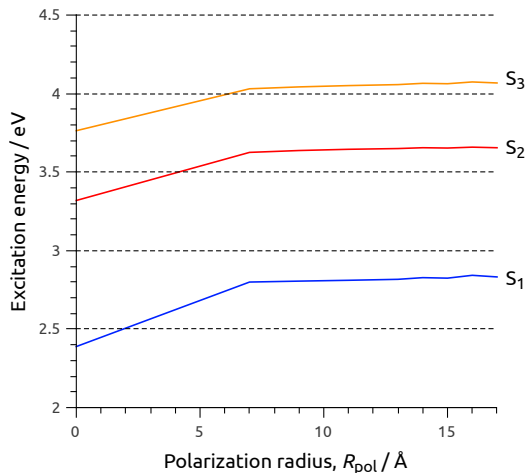the calculations presented.

Figure 3: Lowest three excitation energies of betaine-30 as functions of the polarization radius, R_{pol}, with the total radius for the inclusion of classical environment fixed at 25 Å. Blue, red and yellow curves refer to the first, second and third excited states, respectively, calculated on one snapshot at TDDFT level, within the linear response approach. A value of $R_{\text{pol}} = 0$ indicates that all the solvent molecules are non polarizable, and are only described in terms of fixed multipoles.

All calculations (geometry optimizations and excitation energies) have been performed at the (TD)DFT level of theory using the CAM-B3LYP exchange-correlation functional[53] together with the 6-31+G(d,p) gaussian basis set. We particularly focused on the properties of the lowest (bright) excitation, on which the $E_{\text{T}}(30)$ scale is based. The excitation shows a strong charge-transfer character, and for this reason the corrected linear response (c-LR) approach[28] is expected to give a more accurate picture than the standard LR.

The results obtained with the three different QM/AMOEBA responses described in section 2.3 (namely ω_0, LR and c-LR) are summarized in Figure 4.
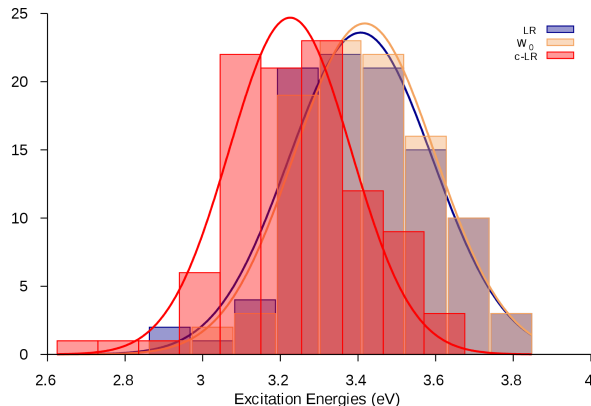
Figure 4: Distributions of the lowest (bright) excitation of betaine-30 calculated with the three different QM/AMOEBA responses (LR: blue line, ω_0: yellow line, c-LR: red line). The curves are gaussian fits of the histograms.

From the comparison, it is evident that including the effects of the relaxation of the solvent response through a c-LR formulation a significant redshift (of about 0.2 eV) is found with respect to the frozen solvent approximation (ω_0). As expected due to the CT character of the excitation, the LR formulation does not change the ω_0 value as it cannot account for the effects of the rearrangement of electron density undergone by the dye upon excitation.

The comparison with experimental data is, however, the most interesting test. In order to be able to provide an analysis which is not biased by systematic errors in the excitation energies due to the QM level of theory, we compare the computed and experimental gas-to-water solvatochromic shifts instead than the absolute excitation energies. To better elucidate the various effects that determine the observed solvatochromic shift, i.e., short-range and specific interactions and the bulk effects, it is useful to compare with two different solvation models. The first, QM/PCM, employs a purely continuum description of the solvent whereas the second, QM(ME)/PCM, includes a "minimal environment" in the QM system, i.e., the two water molecules hydrogen-bonded to the oxygen of the dye, while the rest of the solvent is still treated at PCM level. The latter model should better describe the combination of short-range specific and bulk interaction with respect to the QM/PCM analog. In the QM(ME)/PCM calculation the configuration of the QM water molecules has been optimized

18

at CAM-B3LYP/6-31+G(d,p). All calculations refer to c-LR responses.

Table 1: Gas-to-water solvatochromic shift (in eV) for the lowest (bright) excitation of betaine-30, computed with different solvation models. The QM/AMOEBA value is obtained from the average over 200 snapshots extracted from the MD. The confidence interval is indicated. All calculations were performed at TDCAM-B3LYP/6-31+G(d,p) level.

QM/PCM	0.79
QM(ME)/PCM	1.17
QM/AMOEBA	1.58 ± 0.02
Exp[54]	1.56

As expected, the QM/PCM model markedly underestimates the solvatochromic shift with respect to the two other models (and to experiments). This clearly shows that including an atomistic description of the most strongly interacting water molecules is fundamental to account for the differential solvation effects in the ground and the excited states of the betaine-30. However, it is worth noting that the inclusion of the two hydrogen bonded water molecules in the QM region in combination with a continuum description (through the QM(ME)/PCM model) still fails to recover a large portion of the observed solvatochromic effects (namely 0.4 eV are still missing to reproduce the experiments). The inclusion of a larger number of explicit solvent molecules with AMOEBA instead leads to a solvatochromic shift which is almost exactly equivalent to the experiments. Such an excellent agreement might be fortuitous due to the many approximations introduced in the comparison, such as the use of a calculated vertical excitation and an experimental band maximum, and the extrapolated value used for the experimental gas-phase value. However, the additional +0.4 eV shift obtained going from the QM(ME)/PCM to the QM/AMOEBA model clearly suggests that the strong hydrogen-bonding effect is not the only source of differences with respect to a purely bulk description. Looking more deeply into the results, it is interesting to point out that the difference between QM/AMOEBA and QM(ME)/PCM solvatochromic shifts is already present at ω_0 level, which means that the addition of the c-LR response is almost equally described by the two models. We can thus speculate that the zwitterionic nature of betaine-30 ground state has a large, specific, long range effect in orienting the

water molecules and this leads to a final solvation effect that cannot be accurately described with a pure continuum or a partial atomistic/continuum approach.

4   Conclusions

We presented here the theoretical development and the computational implementation of a polarizable embedding QM/MM within the AMOEBA framework. The implementation accounts for a fully relaxed evaluation of the QM/MM energy both for ground and excited states, in the framework of a DFT/TD-DFT theory. In particular for the excited state, a state-specific (SS) formulation of the response of the classical part of the system has been used following the c-LR approach originally developed within continuum models.[28] The test case of betaine-30 is presented and discussed in comparison with purely continuum and mixed atomistic/continuum models: the obtained results in the simulation of the large gas-to-water solvatochromism of this dye, show that the QM/AMOEBA approach can describe with equal accuracy the effect of water on the zwitterionic ground state and the modification induced by the excitation.

This work represents the first step of a series of new developments towards a high-performance, parallel implementation of polarizable QM/MM molecular dynamics, which stems from the recent developments in the implementation of AMOEBA for classical MD simulations in the newly developed Tinker-HP[34,55–57] code and with the implementation of a versatile and transparent Tinker HP/Gaussian interface for energy, gradients and properties. The implementation of analytical gradients for the QM/AMOEBA model and the development of the Gaussian/AMOEBA interface will allow us to exploit not only parallelism, but also linear scaling techniques[46] in order to further reduce the overall computational cost due to the polarizable embedding. The addition of a further layer to the model, namely, of a polarizable continuum solvation model, will finally allow to deal with long range, bulk effects while, at the same time, reducing the portion of environment to be treated explic-

itly. An innovative, parallel, linear-scaling implementation of the Conductor-like Screening Model[38,58,59] has already been coupled to AMOEBA for classical MD simulations[38] and to the MMpol dipole-based polarizable force field for QM/MM calculations:[46] the combination of all the aforementioned developments will result in a powerful and efficient tool to investigate dynamic properties and reactivity in complex environments.

Acknowledgement

References

(1) Cramer, C.; Truhlar, D. *Chem. Rev.* **1999**, *99*, 2161–2200.

(2) Tomasi, J.; Mennucci, B.; Cammi, R. *Chem. Rev.* **2005**, *105*, 2999–3093.

(3) Warshel, A.; Levitt, M. *J. Mol. Biol.* **1976**, *103*, 227–249.

(4) Gao, J.; Xia, X. *Science* **1992**, *258*, 631–635.

(5) Maseras, F.; Morokuma, K. *J. Comput. Chem.* **1995**, *16*, 1170–1179.

(6) Lin, H.; Truhlar, D. G. *Theor. Chem. Acc.* **2006**, *117*, 185–199.

(7) Senn, H. M.; Thiel, W. *Angew. Chem.* **2009**, *48*, 1198–1229.

(8) Thompson, M. A.; Schenter, G. K. *J. Phys. Chem.* **1995**, *99*, 6374–6386.

(9) Bryce, R. A.; Buesnel, R.; Hillier, I. H.; Burton, N. A. *Chem. Phys. Lett.* **1997**, *279*, 367–371.

(10) Gordon, M. S.; Freitag, M. A.; Bandyopadhyay, P.; Jensen, J. H.; Kairys, V.; Stevens, W. J. *J. Phys. Chem. A* **2001**, *105*, 293–307.

(11) Jensen, L.; van Duijnen, P. T.; Snijders, J. G. *J. Chem. Phys.* **2003**, *119*, 3800–3809.

(12) Nielsen, C. B.; Christiansen, O.; Mikkelsen, K. V.; Kongsted, J. *J. Chem. Phys.* **2007**, *126*, 154112–154118.

(13) Illingworth, C. J. R.; Parkes, K. E. B.; Snell, C. R.; Ferenczy, G. G.; Reynolds, C. A. *J. Phys. Chem. A* **2008**, *112*, 12151–12156.

(14) Curutchet, C.; Muñoz-Losa, A.; Monti, S.; Kongsted, J.; Scholes, G. D.; Mennucci, B. *J. Chem. Theory Comput.* **2009**, *5*, 1838–1848.

(15) Lipparini, F.; Barone, V. *J. Chem. Theory Comput.* **2011**, *7*, 3711–3724.

(16) Steindal, A. H.; Ruud, K.; Frediani, L.; Aidas, K.; Kongsted, J. *J. Phys. Chem. B* **2011**, *115*, 3027–3037.

(17) Boulanger, E.; Thiel, W. *J. Chem. Theory Comput.* **2012**, *8*, 4527–4538.

(18) Gordon, M. S.; Fedorov, D. G.; Pruitt, S. R.; Slipchenko, L. V. *Chem. Rev.* **2012**, *112*, 632–672.

(19) Gao, J.; Truhlar, D. G.; Wang, Y.; Mazack, M. J. M.; Löffler, P.; Provorse, M. R.; Rehak, P. *Acc. Chem. Res.* **2014**, *47*, 2837–2845.

(20) Zeng, Q.; Liang, W. *J. Chem. Phys.* **2015**, *143*, 134104–134119.

(21) Ren, P.; Ponder, J. W. *J. Phys. Chem. B* **2003**, *107*, 5933–5947.

(22) Grossfield, A.; Ren, P.; Ponder, J. W. *J. Am. Chem. Soc.* **2003**, *125*, 15671–15682.

(23) Ponder, J. W.; Wu, C.; Ren, P.; Pande, V. S.; Chodera, J. D.; Schnieders, M. J.; Haque, I.; Mobley, D. L.; Lambrecht, D. S.; DiStasio Jr, R. A.; Head-Gordon, M.; Clark, G. N. I.; Johnson, M. E.; Head-Gordon, T. *J. Phys. Chem. B* **2010**, *114*, 2549–2564.

(24) Kratz, E. G.; Walker, A. R.; Lagardère, L.; Lipparini, F.; Piquemal, J.-P.; Andrés Cisneros, G. *J. Comp. Chem.* **2016**, *37*, 1019–1029.

(25) Corni, S.; Cammi, R.; Mennucci, B.; Tomasi, J. *J. Chem. Phys.* **2005**, *123*, 134512–134522.

(26) Cupellini, L.; Amovilli, C.; Mennucci, B. *J. Phys. Chem. B* **2015**, *119*, 8984–8991.

(27) Guido, C. A.; Jacquemin, D.; Adamo, C.; Mennucci, B. *J. Chem. Theory Comput.* **2015**, *11*, 5782–5790.

(28) Caricato, M.; Mennucci, B.; Tomasi, J.; Ingrosso, F.; Cammi, R.; Corni, S.; Scalmani, G. *J. Chem. Phys.* **2006**, *124*, 124520–124533.

(29) Improta, R.; Barone, V.; Scalmani, G.; Frisch, M. J. *J. Chem. Phys.* **2006**, *125*, 054103–054112.

(30) Marenich, A. V.; Cramer, C. J.; Truhlar, D. G.; Guido, C. A.; Mennucci, B.; Scalmani, G.; Frisch, M. J. *Chem. Sci.* **2011**, *2*, 2143–2161.

(31) Reichardt, C. *Chem. Rev.* **1994**, *94*, 2319–2358.

(32) Machado, V. G.; Stock, R. I.; Reichardt, C. *Chem. Rev.* **2014**, *114*, 10429–10475.

(33) Thole, B. T. *Chem. Phys.* **1981**, *59*, 341–350.

(34) Lipparini, F.; Lagardère, L.; Stamm, B.; Cancs, E.; Schnieders, M.; Ren, P.; Maday, Y.; Piquemal, J.-P. *J. Chem. Theory Comput.* **2014**, *10*, 1638–1651.

(35) van Duijnen, P. T.; Swart, M. *J. Phys. Chem. A* **1998**, *102*, 2399–2407.

(36) Wang, J.; Cieplak, P.; Li, J.; Wang, J.; Cai, Q.; Hsieh, M.; Lei, H.; Luo, R.; Duan, Y. *J. Phys. Chem. B* **2011**, *115*, 3100–3111.

(37) Sala, J.; Guàrdia, E.; Masia, M. *J. Chem. Phys.* **2010**, *133*, 234101–234115.

(38) Lipparini, F.; Lagardère, L.; Raynaud, C.; Stamm, B.; Cancès, E.; Mennucci, B.; Schnieders, M.; Ren, P.; Maday, Y.; Piquemal, J. P. *J. Chem. Theory Comput.* **2015**, *11*, 623–634.

(39) Lipparini, F.; Scalmani, G.; Mennucci, B.; Cancès, E.; Caricato, M.; Frisch, M. J. *J. Chem. Phys.* **2010**, *133*, 014106–014117.

(40) Lipparini, F.; Scalmani, G.; Mennucci, B.; Frisch, M. J. *J. Chem. Theory Comput.* **2011**, *7*, 610–617.

(41) Lipparini, F.; Cappelli, C.; Barone, V. *J. Chem. Theory Comput.* **2012**, *8*, 4153–4165.

(42) Casida, M. E.; Jamorski, C.; Casida, K. C.; Salahub, D. R. *J. Chem. Phys.* **1998**, *108*, 4439–4449.

(43) Cammi, R.; Corni, S.; Mennucci, B.; Tomasi, J. *J. Chem. Phys.* **2005**, *122*, 104513–104525.

(44) Handy, N. C.; Schaefer, H. F. *J. Chem. Phys.* **1984**, *81*, 5031–5033.

(45) 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, Jr., J. A.; 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, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian 09 Revision A.1. 2009; `www.gaussian.com`, Gaussian, Inc., Wallingford CT.

(46) Caprasecca, S.; Jurinovich, S.; Lagardère, L.; Stamm, B.; Lipparini, F. *J. Chem. Theory Comput.* **2015**, *11*, 694–704.

(47) Greengard, L.; Rokhlin, V. *J. Comput. Phys.* **1987**, *73*, 325–348.

(48) Cancès, E.; Mennucci, B.; Tomasi, J. *J. Chem. Phys.* **1997**, *107*, 3032–3041.

(49) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926–935.

(50) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. *J. Comput. Phys.* **1977**, *23*, 327–341.

(51) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684–3690.

(52) Case, D. A.; Babin, V.; Berryman, J. T.; Betz, R. M.; Cai, Q.; Cerutti, D. S.; Cheatham III, T. E.; Darden, T. A.; Duke, R. E.; Gohlke, H.; Goetz, A. W.; Gusarov, S.; Homeyer, N.; Janowski, P.; Kaus, J.; Kolossvry, I.; Kovalenko, A.; Lee, T. S.; LeGrand, S.; Luchko, T.; Luo, R.; Madej, B.; Merz, K. M.; Paesani, F.; Roe, D. R.; Roitberg, A.; Sagui, C.; Salomon-Ferrer, R.; Seabra, G.; Simmerling, C. L.; Smith, W.; Swails, J.; Walker, R. C.; Wang, J.; Wolf, R. M.; Wu, X.; A, K. P. AMBER 14. 2014; `http://ambermd.org/`, University of California, San Francisco.

(53) Yanai, T.; Tew, D. P.; Handy, N. C. *Chem. Phys. Lett.* **2004**, *393*, 51–57.

(54) Reichardt, C.; Welton, T. *Solvents and Solvent Effects in Organic Chemistry*, 4th ed.; Wiley-VCH: Weinheim, Germany, 2011.

(55) Lagardère, L.; Lipparini, F.; Stamm, B.; Polack, E.; Jolly, L.; Narth, C.; Kratz, E.; Cisneros, G.; Schnieders, M.; Darden, T.; Gresh, N.; Maday, Y.; Ponder, J.; Ren, P.; Piquemal, J.-P. TINKER-HP. Sorbonne Universitè, UPMC Univ. Paris 06, 2016.

(56) Lagardère, L.; Lipparini, F.; Polack, E.; Stamm, B.; Cancès, E.; Schnieders, M.; Ren, P.; Maday, Y.; Piquemal, J.-P. *J. Chem. Theory Comput.* **2015**, *11*, 2589–2599.

(57) Narth, C.; Lagardère, L.; Polack, E.; Gresh, N.; Wang, Q.; Bell, D. R.; Rackers, J. A.; W., P. J.; Ren, P.; Piquemal, J.-P. *J. Comput. Chem.* **2016**, *37*, 494–505.

(58) Klamt, A.; Schuurmann, G. *J. Chem. Soc., Perkin Trans. 2* **1993**, 799–805.

(59) Lipparini, F.; Stamm, B.; Cancès, E.; Maday, Y.; Mennucci, B. *J. Chem. Theory Comput.* **2013**, *9*, 3637–3648.