

Electron dynamics in complex environments with real-time time dependent density functional theory in a QM-MM framework

Uriel N. Morzan, Francisco F. Ramírez, M. Belén Oviedo, Cristián G. Sánchez, Damián A. Scherlis, and Mariano C. González Lebrero

Citation: *The Journal of Chemical Physics* **140**, 164105 (2014); doi: 10.1063/1.4871688

View online: <http://dx.doi.org/10.1063/1.4871688>

View Table of Contents: <http://scitation.aip.org/content/aip/journal/jcp/140/16?ver=pdfcov>

Published by the [AIP Publishing](#)

Articles you may be interested in

[Quantum mechanical/molecular mechanical/continuum style solvation model: Time-dependent density functional theory](#)

J. Chem. Phys. **139**, 084106 (2013); 10.1063/1.4819139

[Optical properties of an organic dye from time-dependent density functional theory with explicit solvent: The case of alizarin](#)

J. Chem. Phys. **137**, 154314 (2012); 10.1063/1.4758877

[Efficient electron dynamics with the planewave-based real-time time-dependent density functional theory: Absorption spectra, vibronic electronic spectra, and coupled electron-nucleus dynamics](#)

J. Chem. Phys. **135**, 244112 (2011); 10.1063/1.3671952

[Solvent effects on optical properties of molecules: A combined time-dependent density functional theory/effective fragment potential approach](#)

J. Chem. Phys. **129**, 144112 (2008); 10.1063/1.2992049

[Assignment of the complex vibrational spectra of the hydrogenated ZnO polar surfaces using QM/MM embedding](#)

J. Chem. Phys. **118**, 317 (2003); 10.1063/1.1523897



Re-register for Table of Content Alerts

Create a profile.



Sign up today!



Electron dynamics in complex environments with real-time time dependent density functional theory in a QM-MM framework

Uriel N. Morzan,¹ Francisco F. Ramírez,¹ M. Belén Oviedo,² Cristián G. Sánchez,² Damián A. Scherlis,^{1,a)} and Mariano C. González Lebrero^{3,a)}

¹*Departamento de Química Inorgánica, Analítica y Química Física/INQUIMAE, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Ciudad Universitaria, Pab. II, Buenos Aires (C1428EHA), Argentina*

²*Departamento de Matemática y Física, Facultad de Ciencias Químicas, INFIQC, Universidad Nacional de Córdoba, Ciudad Universitaria, X5000HUA Córdoba, Argentina*

³*Instituto de Química y Físicoquímica Biológicas, IQUIFIB, CONICET, Argentina*

(Received 16 January 2014; accepted 7 April 2014; published online 25 April 2014)

This article presents a time dependent density functional theory (TDDFT) implementation to propagate the Kohn-Sham equations in real time, including the effects of a molecular environment through a Quantum-Mechanics Molecular-Mechanics (QM-MM) hamiltonian. The code delivers an all-electron description employing Gaussian basis functions, and incorporates the Amber force-field in the QM-MM treatment. The most expensive parts of the computation, comprising the commutators between the hamiltonian and the density matrix—required to propagate the electron dynamics—, and the evaluation of the exchange-correlation energy, were migrated to the CUDA platform to run on graphics processing units, which remarkably accelerates the performance of the code. The method was validated by reproducing linear-response TDDFT results for the absorption spectra of several molecular species. Two different schemes were tested to propagate the quantum dynamics: (i) a leap-frog Verlet algorithm, and (ii) the Magnus expansion to first-order. These two approaches were confronted, to find that the Magnus scheme is more efficient by a factor of six in small molecules. Interestingly, the presence of iron was found to seriously limitate the length of the integration time step, due to the high frequencies associated with the core-electrons. This highlights the importance of pseudopotentials to alleviate the cost of the propagation of the inner states when heavy nuclei are present. Finally, the methodology was applied to investigate the shifts induced by the chemical environment on the most intense UV absorption bands of two model systems of general relevance: the formamide molecule in water solution, and the carboxy-heme group in Flavohemoglobin. In both cases, shifts of several nanometers are observed, consistently with the available experimental data.

© 2014 AIP Publishing LLC. [<http://dx.doi.org/10.1063/1.4871688>]

I. INTRODUCTION

The interest in the calculation of excited states and spectroscopic transitions has motivated in the last 20 years a pronounced expansion in the application of time dependent density functional theory (TDDFT).^{1,2} This methodology has become the first choice to compute these properties in middle-sized molecular and extended systems, for which the much more expensive quantum-chemistry schemes based on multi-configurational wavefunctions used to be the standard approach. Presently, the vast majority of applications of TDDFT makes use of the linear response formulation (LR-TDDFT). In this framework, the perturbation of the Kohn-Sham density by an applied field is expressed to linear order introducing the susceptibility (or density-density response function) of the non-interacting electron system.^{3,4} Excitation energies and related quantities can be obtained from the poles of the susceptibility function. Different treatments have been devised to implement this technique; in particular, the Casida formulation⁵ is the one adopted by most LR-TDDFT quantum-chemistry

programs to compute excitation frequencies and oscillator strengths.

Alternatively, it is possible to explicitly evolve the time-dependent Kohn-Sham equations to simulate the electron dynamics resulting from a perturbation. This leads to the so-called real-time time dependent density functional theory (RT-TDDFT), which in principle provides the full temporal description of the Kohn-Sham single-particle states. The spectroscopic frequencies and intensities can be recovered from the Fourier transform of the time-dependent dipole moment, but in general this requires to simulate the electron dynamics over some tens of femtoseconds. This is computationally costly and explains why the real-time formulation of TDDFT has not been of widespread application as its linear response counterpart. However, RT-TDDFT reproduces the quantum-dynamics of the electrons, and therefore may provide insights on fundamental phenomena banned to LR-TDDFT: charge transport, photocurrent generation, coupled ion-electron dynamics, or the response to strong (laser) fields beyond the linear regime.

In parallel with the development of LR-TDDFT codes, a smaller number of real time implementations using different

^{a)} Authors to whom correspondence should be addressed. Electronic mails: damian@qi.fcen.uba.ar and mcgl@qb.ffyb.uba.ar

basis functions and numerical schemes have appeared and evolved since 1990s. The earliest versions employed real space grids and the local density approximation, for example, the pioneering work by Bertsch and Yabana,⁶ which is the antecessor of the Octopus TDDFT package.⁷ Other kind of basis sets were adopted in different developments, including plane-waves,^{8,9} numerical orbitals (in the context of the SIESTA program),¹⁰ or, more recently, Gaussian functions.^{11–13} Real-time time dependent DFT simulations were also carried out at the semiempirical level, using the DFTB method.^{14–16} Furthermore, first-principles ion-electron dynamics were performed based on such implementations, applying the Ehrenfest scheme or alternative treatments.^{8,16}

In this article, we introduce a real time implementation of TDDFT using Gaussian basis sets with two particular features: (i) it is coupled to a Quantum-Mechanics Molecular-Mechanics scheme which allows for the representation of complex environments, and (ii) it is significantly accelerated by performing part of the computations on graphics processing units (GPU). Hybrid quantum-mechanics molecular-mechanics (QM-MM) techniques introduce self-consistently a set of atomic charges from an empirical force-field in the *ab initio* hamiltonian. In this way, the electronic density of the “quantum” subsystem is modulated by the presence of a non-reactive environment, for example, a solute surrounded by the solvent molecules, or an active site immersed in a protein matrix. The embedding of RT-TDDFT in a QM-MM framework is alluring for various applications involving large systems where the environment does not directly participate in the quantum-dynamics, as for example, spectroscopy in solution, or electron transport across solid-liquid interfaces, or between redox centers in metallo-proteins.

In Sec. II, the methodology is described in detail. Next, the implementation is validated by confronting the spectroscopic data obtained for simple molecules, against the results obtained through linear response TDDFT. Efficiency aspects are examined in Sec. III B, and finally, the last two parts of the paper (Secs. III C and III D) are devoted to investigate the role of the environment on the spectral features of two systems of chemical and biological relevance: (i) the solvatochromic effect of water on the position of the intense $\pi_{nb} \rightarrow \pi^*$ transition characteristic of the peptidic bond, and (ii) the location of the Soret band in a heme-protein.

II. METHODOLOGICAL IMPLEMENTATION

A. DFT and QM-MM scheme

The present real-time implementation is based on an all-electron, Gaussian basis sets density functional code developed in our group. In this code, the exchange correlation energy is computed numerically in a real space grid, and represents the most expensive part of the self-consistent procedure. Recently, we have reformulated this grid to an atom centered radial mesh where the density is grouped and partitioned into smaller points, and articulated it in CUDA to perform the integration of the exchange-correlation potential in graphics processing units.¹⁷ With this migration of architecture, the SCF procedure has been accelerated by a factor

of 20 when running on a GTX 780 GPU, with respect to the time consumed in a single CPU processor.¹⁷ Furthermore, this code has been interfaced with a classical force-field, to carry out hybrid quantum-mechanics molecular-mechanics simulations. In such schemes, the system is partitioned in two regions: the “quantum” or QM subsystem, where the electronic or chemical processes of interest take place, and the “classical” or MM region, containing the environment represented by a distribution of atomic charges which interacts electrostatically with the quantum part of the system. Additionally, the atoms of one subsystem may interact with those of the other through Lennard-Jones or any other empirical potential. The total energy E_{tot} takes the form:

$$E_{tot} = E_{QM} + E_{MM} + E_{QM-MM}, \quad (1)$$

where E_{QM-MM} and E_{MM} denote the energy of the quantum and classical parts, computed as usual, and E_{QM-MM} is the hybrid contribution:

$$E_{QM-MM} = \sum_A q_A \int \frac{d\mathbf{r}\rho(\mathbf{r})}{\mathbf{r} - \mathbf{r}_A} + \sum_{A,B} \frac{q_A Z_B}{\mathbf{r}_A - \mathbf{r}_B} + E_{QM-MM}^{LJ}. \quad (2)$$

In the expression above, $\rho(\mathbf{r})$ is the electron density, q_A represents the point charges of the MM atoms, and Z_A the nuclear charges of the QM atoms (with the indices A and B running over the atoms of the classical and the quantum regions, respectively). E_{QM-MM}^{LJ} introduces parameterized, two-body interactions between the atoms of both subsystems. This term plays an important role in geometry optimizations or molecular dynamics, because it prevents that negative point charges from the MM region collapse onto the positive QM nuclei. Additional terms may be present if the QM and MM partitions are connected via chemical bonds. All these contributions are included in the hamiltonian across the SCF cycle, so that the electron density adapts self-consistently to the point charge distribution. In the same way, the forces on the MM region are subject to the effect of the electron density, and therefore it is said that each half of the system responds self-consistently to the other. In the implementation presented in this article, the QM region is treated via the Kohn-Sham hamiltonian, whereas the MM atoms are described by the AMBER force-field.^{17,18}

B. Temporal evolution of the electron density

The Liouville-von Neumann equation describes the electron dynamics of a quantum system in terms of the density matrix \mathbf{P} :¹⁹

$$i\hbar \frac{\partial \mathbf{P}}{\partial t} = [\mathbf{H}, \mathbf{P}] \quad (3)$$

where \mathbf{H} is the electronic hamiltonian matrix. In the present implementation \mathbf{H} is the Kohn-Sham matrix and \mathbf{P} is constructed from the coefficients $C_{\mu a}$ of the canonical (orthonormal) molecular orbitals ψ_a :

$$\psi_a(\mathbf{r}, t) = \sum_{\mu} C_{\mu a}(t) \phi_{\mu}(\mathbf{r}), \quad (4)$$

$$P_{\nu\mu} = \sum_a C_{\mu a}(t) C_{\nu a}^*(t) \quad (5)$$

with ϕ_μ the Gaussian basis functions. At time $t = 0$ a perturbation V^{app} is applied and, starting from the ground state, the density matrix is propagated using Eq. (3) in a discretized temporal domain. Two different techniques were tested to integrate this equation, which are described in the next subsection. Once \mathbf{P} has been evolved for the next time step Δt , the Kohn-Sham matrix needs to be updated because it depends on the density matrix via the Coulomb (\mathbf{V}^c) and the exchange-correlation (\mathbf{V}^{xc}) terms:

$$\mathbf{H}^{AO}(t) = \mathbf{H}^{core} + \mathbf{V}^c[\mathbf{P}(t)] + \mathbf{V}^{xc}[\mathbf{P}(t)] + \mathbf{V}_{app}(t), \quad (6)$$

$$V_{\mu\nu}^c(t) = \sum_{\lambda\sigma} P_{\lambda\sigma}^{AO}(t) \int d\mathbf{r}_1 d\mathbf{r}_2 \frac{\phi_\mu(\mathbf{r}_1)\phi_\nu(\mathbf{r}_1)\phi_\sigma(\mathbf{r}_2)\phi_\lambda(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|}, \quad (7)$$

$$V_{\mu\nu}^{xc}(t) = \int v_{xc}[\rho(\mathbf{r}, t)]\phi_\mu(\mathbf{r})\phi_\nu(\mathbf{r})d\mathbf{r}, \quad (8)$$

with v_{xc} the exchange correlation potential. The supraindex *AO* highlights that Eq. (7) is valid for \mathbf{P} expressed in the basis of (Gaussian) atomic orbitals and that $\mathbf{H}^{AO}(t)$ is assembled in that basis. Therefore, the density matrix needs to be transformed to the atomic basis before evaluating \mathbf{H} :

$$\mathbf{P}^{AO}(t) = \mathbf{X}\mathbf{P}(t)\mathbf{X}^\dagger. \quad (9)$$

Here \mathbf{X} is the transformation matrix that diagonalizes the overlap matrix \mathbf{S} ,

$$\mathbf{X}^\dagger\mathbf{S}\mathbf{X} = \mathbf{1} \quad S_{\mu\nu} = \int \phi_\mu(\mathbf{r})\phi_\nu(\mathbf{r})d\mathbf{r},$$

which in the present implementation is calculated using the DSPEV routine of the LAPACK library. Naturally, once the Kohn-Sham matrix is obtained, it has to be transformed to the molecular orbitals (MO) basis to go ahead with the integration of Eq. (3):

$$\mathbf{H}(t) = \mathbf{X}^\dagger\mathbf{H}^{AO}(t)\mathbf{X}. \quad (10)$$

The next sequence summarizes the procedure to propagate the quantum dynamics starting from the ground state.

1. Do a standard SCF cycle, from which $\mathbf{H}(t = 0)$ and $\mathbf{P}(t = 0)$ are computed in the MO basis.
2. Compute $[\mathbf{H}(\mathbf{t}), \mathbf{P}(\mathbf{t})]$ to obtain $\mathbf{P}(t + \Delta t)$ via Eq. (3) in one of the ways described in Sec. II C.
3. Transform $\mathbf{P}(t + \Delta t)$ to the AO basis, Eq. (9).
4. Construct $\mathbf{H}(t + \Delta t)$ in the AO basis, Eq. (8).
5. Transform $\mathbf{H}(t + \Delta t)$ to the MO basis, Eq. (10).
6. Go to step 2.

C. Time integration

We examined the performance of two distinct numerical approaches to evolve the electronic equations of motion: the Magnus propagator, and the much simpler Verlet algorithm to first-order. The former approach has been often adopted in real-time quantum dynamics implementations,^{13,16} while the later has been tried successfully in semiempirical RT-TDDFT,¹⁵ but, to the best of our knowledge, no application in first-principles electron dynamics has been reported.

Verlet propagation. A first-order Taylor expansion of $\mathbf{P}(t)$ around t yields:

$$\mathbf{P}(t + \Delta t) = \mathbf{P}(t) + \frac{1}{i\hbar}[\mathbf{H}(t), \mathbf{P}(t)]\Delta t \quad (11)$$

$$\mathbf{P}(t - \Delta t) = \mathbf{P}(t) - \frac{1}{i\hbar}[\mathbf{H}(t), \mathbf{P}(t)]\Delta t$$

The combination of these two equations gives

$$\mathbf{P}(t + \Delta t) = 2[\mathbf{H}(t), \mathbf{P}(t)]\Delta t + \mathbf{P}(t - \Delta t) \quad (12)$$

so that the propagation of the density matrix requires its value at the present and the previous time step, and also the value of the Kohn-Sham matrix at the current step. We note that this propagation scheme preserves by construction the trace of the density matrix, because the trace is invariant against the permutation of the matrix product,

$$Tr(\mathbf{H}\mathbf{P}) = \sum_j \sum_i H_{ji} P_{ij}$$

$$Tr(\mathbf{P}\mathbf{H}) = \sum_j \sum_i P_{ji} H_{ij} = \sum_i \sum_j P_{ij} H_{ji} = Tr(\mathbf{H}\mathbf{P}).$$

and therefore the trace of the commutator in Eq. (12) is zero.

Magnus expansion. Given a time-independent hamiltonian matrix \mathbf{H} , the differential Eq. (3) has the following exact solution:

$$\mathbf{P}(t) = e^{-i\mathbf{H}t/\hbar}\mathbf{P}(0)e^{i\mathbf{H}t/\hbar}. \quad (13)$$

To see this, it suffices to differentiate the right-hand side of Eq. (13), and to note that the exponential operator commutes with \mathbf{H} . In the case of a time-dependent hamiltonian $\mathbf{H}(t)$, an equivalent solution would be

$$\mathbf{P}(t) = e^{-i\int_0^t \mathbf{H}(t)dt/\hbar}\mathbf{P}(0)e^{i\int_0^t \mathbf{H}(t)dt/\hbar}. \quad (14)$$

However, a Taylor expansion shows that this expression leads to an inconsistency in the chronological order in which $\mathbf{H}(t)$ operates (see Ref. 19). Magnus worked this out demonstrating that the propagator could be reformulated as follows:^{19,20}

$$\mathbf{P}(t) = e^{\mathbf{A}(t)}\mathbf{P}(0)e^{-\mathbf{A}(t)}\mathbf{A}(t) = \mathbf{A}_1(t) + \mathbf{A}_2(t) + \mathbf{A}_3(t) + \dots \quad (15)$$

$$\mathbf{A}_1(t) = \frac{1}{i\hbar} \int_0^t dt_1 \mathbf{H}(t_1),$$

$$\mathbf{A}_2(t) = -\frac{1}{2} \left(\frac{1}{i\hbar}\right)^2 \int_0^t dt_2 \int_0^{t_2} dt_1 [\mathbf{H}(t_1), \mathbf{H}(t_2)],$$

$$\mathbf{A}_3(t) = -\frac{1}{6} \left(\frac{1}{i\hbar}\right)^3 \int_0^t dt_3 \int_0^{t_3} dt_2 \int_0^{t_2} dt_1 [\mathbf{H}(t_1), [\mathbf{H}(t_2), \mathbf{H}(t_3)]] + [[\mathbf{H}(t_1), \mathbf{H}(t_2)], \mathbf{H}(t_3)].$$

Following recent implementations,^{13,16} we adopt here the Magnus expansion to first-order—including only $\mathbf{A}_1(t)$ in equation (15)—, the integral by a middle-point approximation, and apply the Baker-Campbell-Hausdorff (BCH) formula to calculate the matrix product with the exponential

term:

$$\mathbf{A}_1(t) \approx \frac{1}{i\hbar} \mathbf{H}(t/2)t \quad (16)$$

$$\begin{aligned} \mathbf{P}(t) = & \mathbf{P}(0) - it[\mathbf{A}_1(t), \mathbf{P}(0)] - \frac{t^2}{2!}[\mathbf{A}_1(t), [\mathbf{A}_1(t), \mathbf{P}(0)]] \\ & + i \frac{t^3}{3!}[\mathbf{A}_1(t), [\mathbf{A}_1(t), [\mathbf{A}_1(t), \mathbf{P}(0)]]] + \dots \quad (17) \end{aligned}$$

A caveat of this approach is that the value of $\mathbf{H}(t/2)$ in Eq. (16) depends on the (unknown) density matrix at time $t/2$. To deal with this, we have adopted the predictor-corrector scheme proposed by Van Voorhis and co-workers,¹¹ where $\mathbf{H}(t/4)$ is estimated from a linear interpolation based on the two preceding matrices $\mathbf{H}(-t/2)$ and $\mathbf{H}(-3t/4)$. This allows to compute $\mathbf{A}_1(t/2)$, which is used to propagate the density matrix half-step. The resulting $\mathbf{P}(t/2)$ is employed to construct the Kohn-Sham matrix $\mathbf{H}(t/2)$, with which the dynamics is evolved a full time step. This procedure is correct to order Δt^2 .¹¹ The effect of truncating expansion (17) has been investigated and is discussed in Sec. III B.

D. Absorption spectra

An optical absorption spectrum can be obtained from the response of the dipole moment to an electric perturbation in a real time simulation. This can be achieved by introducing a δ function like electric-field kick capable to excite all electronic frequencies with the same intensity.⁶ In this work, a linearly polarized Gaussian type perturbation $\vec{E}(t)$ was imposed to the ground state density to simulate a classical electric field in the linear regime:

$$\vec{E}(t) = E_0 \exp[-(t - t_0)^2/2\omega^2] \hat{r}, \quad (18)$$

where t_0 is the time for which the pulse reaches its maximum value E_0 , ω is its width, and \hat{r} is the unitary vector pointing in the direction of x , y or z . Typical values used in this work for E_0 and ω are in the order of 0.05 and 2.23 atomic units, respectively. The coupling term between the external (applied) electric field and the system is given by

$$V_{app}(t) = -\vec{E}(t) \bullet \vec{\mu}(t), \quad (19)$$

where $\vec{\mu}(t)$ is the time dependent dipole moment, which scalar components can be approximated to first order by

$$\mu_k(t) = \mu_k^o + \int dt_1 \alpha_{ik}(t - t_1) E_i(t_1). \quad (20)$$

Here μ_k^o is the \hat{k} component of the dipole moment of the system in the absence of an external electric field, and α_{ik} is an element of the linear polarizability tensor. Then, if we define the induced dipole moment μ_k^{ind} as

$$\mu_k^{ind}(t) = \mu_k(t) - \mu_k^o \quad (21)$$

we have an expression for its frequency domain form:

$$\mu_k^{ind}(\omega) = \alpha_{ik}(\omega) E_i(\omega). \quad (22)$$

Finally, the polarizability tensor can be computed from the time-dependent dipole moment and the applied field in the following way:

$$\alpha_{ik}(\omega) = \frac{\mu_k^{ind}(\omega)}{E_i(\omega)} = \frac{\int dt e^{i\omega t} \mu_k^{ind}(t) e^{-\gamma t}}{\int dt e^{i\omega t} E_i(t)}. \quad (23)$$

In the expression above, a damping factor γ was introduced, which produces a broadening in the absorption spectrum peaks that is useful to emulate coupling effects between electronic and nuclear degrees of freedom.^{21,22} Values in the order of 0.2 fs⁻¹ were adopted for this parameter to simulate the spectra. Once the polarizability tensor is obtained in the frequency domain, the absorption cross section tensor $\sigma(\omega)$ and the dipole strength function $S(\omega)$ can be calculated as

$$\sigma(\omega) = \frac{4\pi\omega}{c} Im[\alpha(\omega)], \quad (24)$$

$$S(\omega) = \frac{1}{3} Tr[\sigma(\omega)]. \quad (25)$$

III. APPLICATIONS AND RESULTS

A. Validation: Electronic spectra of simple molecules

To check the reliability of the quantum dynamics produced by this implementation, we have computed the absorption spectra of several molecules following the protocol described in Sec. II D. The list includes a variety of compounds covering an ample range of sizes, from diatomics to tryptophan to carotene (see Figure 1). Simulations times were extended for at least 20 fs, and were performed using the PBE approximation to the exchange-correlation energy,²³ in combination with a DZVP basis set.²⁴ For comparison, electronic excitations were calculated for these systems with linear-response TDDFT, employing the same basis sets and exchange-correlation functional. The Orca²⁵ and the

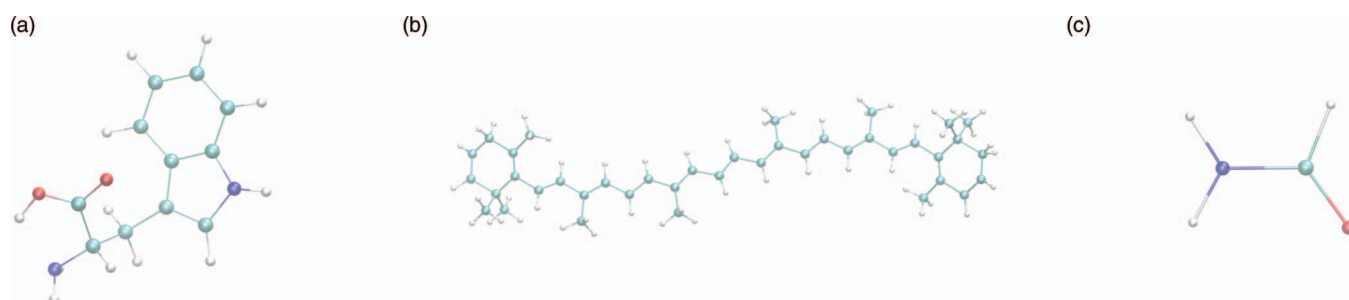


FIG. 1. Some of the molecules investigated in this study: (a) tryptophan, (b) carotene, and (c) formamide. Color code: light blue = carbon; blue = nitrogen; red = oxygen; white = hydrogen.

TABLE I. Calculated and experimental (when available) absorption wavelengths in nm for various molecules in the gas phase. The reported values correspond to the lowest energy transitions. Oscillator strengths are given in parenthesis. Real-time and linear-response TDDFT calculations were performed with the PBE exchange-correlation functional and DZVP basis sets.

	RT-TDDFT	LR-TDDFT	Expt.
CO	147 (0.1)	143 (0.1)	145
H ₂ O	161 (0.04)	159 (0.02)	170
C ₆ H ₆	175 (0.8)	172 (0.6)	180
CH ₄	103 (0.1)	103 (0.1)	127
H ₂	94 (0.3)	95 (0.6)	110
H ₂ SO	224 (0.01)	222 (0.01)	...
Tryptophan	298 (0.1)	311 (0.1)	286
Carotene	568 (0.9)	579 (1.2)	...

Gaussian 03²⁶ packages were used to realize the LR-TDDFT calculations. The comparison between methods and with experiments is presented in Table I. It is seen that the frequencies from real-time simulations are in very good agreement with the linear-response values. The seemingly stronger discrepancies in wavelengths between the LR and RT results for carotene or tryptophan, of around 10 nm, are only apparent: in terms of energy, they represent less than 0.2 eV, i.e., the same discrepancies observed for the smaller molecules absorbing at shorter wavelengths. The LR-TDDFT calculation for carotene exhibits a peak at 545 nm with an oscillator strength of 0.44. The simulation time in RT-TDDFT may not be enough to resolve this peak from the one at 579 nm, entailing a separation of only 0.1 eV, and both would appear as a single signal.

Overall, the computed results are pretty close to the experimental data, with an average error below 15 nm. The intensities provided by the RT-TDDFT approach are also in line with the oscillator strengths calculated with Gaussian 03. To illustrate this, the spectra for both formamide (HCONH₂) and the intermediary species H₂SO are shown in Figure 2, together with the data obtained from LR-TDDFT.

B. Propagators and performance considerations

The Verlet algorithm described above is very inexpensive computationally in comparison with the Magnus or with other conventional quantum-mechanical propagators, but in turn it substantially restrains the length of the time step Δt that can be used to integrate the equations of motion. On the other hand, when applying Magnus propagation to first-order, this time step presumably depends on the number of commutators (NC) included in expansion (17). We examine in this section how these aspects affect the overall performance of the code, in order to identify the most efficient propagation scheme among those implemented.

Table II shows the effect of NC on the maximum time step able to produce a stable, conservative quantum-dynamics. Results are shown for four systems of different size and nature: methane, carotene, the heme group, and the iron dimer. As expected, the length of the time step can be increased if more commutators are considered in expansion (17). For methane and carotene, this effect seems to saturate beyond 15 commutators. It seems perhaps surprising that the

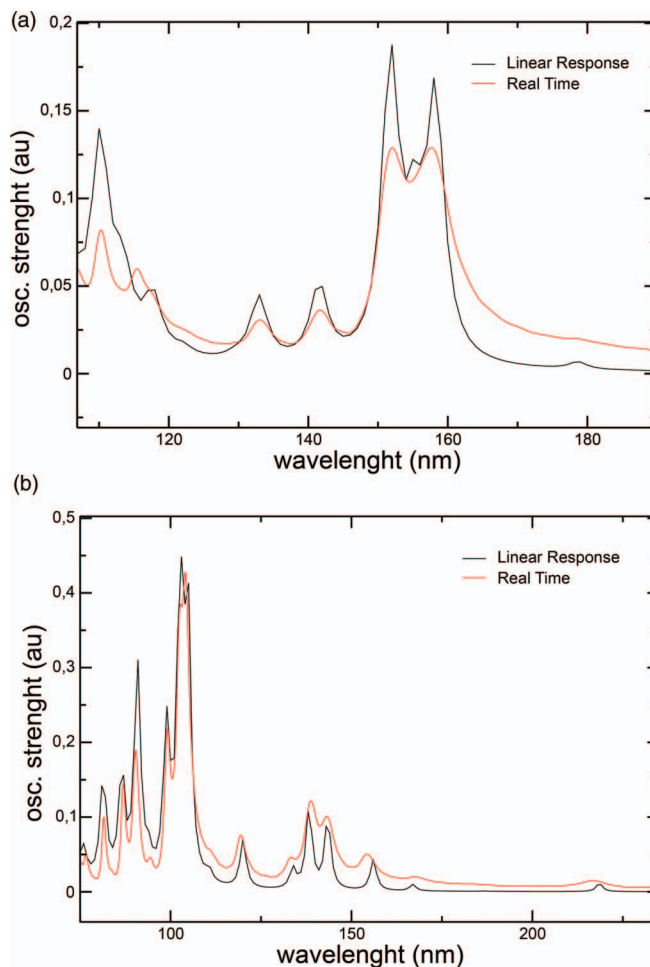


FIG. 2. Simulated spectra for HCONH₂ (a) and H₂SO (b), comparing linear response and real time TDDFT results. Calculations were performed with the PBE exchange-correlation functional and DZVP basis sets.

iron-porphyrin and the Fe₂ species require a sensibly smaller time step than the one used in other systems. The origin of this requirement is not related to the valence *d* shell, but it has to do with the core electrons of the iron atom (we recall that this is an all-electron treatment). This was corroborated through calculations in charged iron dimers lacking the full valence charge density, e.g., Fe₂¹⁸⁺, for which the time step turned out to be exactly the same as in the neutral dimer. Seemingly, the inner electrons subject to the field of a heavy nucleus have much higher frequencies demanding a more stringent integration step. This pinpoints the relevance that pseudopotentials may have when dealing with heavy atoms in RT-TDDFT, not so much because of a reduction of the dimensions of the

TABLE II. The effect of the number of commutators (NC) in the Magnus expansion (Eq. (17)) on the maximum viable time step Δt to propagate the quantum-dynamics a single step. Times are given in femtoseconds.

	NC = 5	NC = 10	NC = 15	NC = 30
Methane	0.002	0.007	0.010	0.012
Carotene	0.005	0.007	0.010	0.010
Fe dimer	0.0003	0.0005	0.0006	0.0009
Heme-group	0.0003	0.0005	0.0006	0.0009

TABLE III. Time step Δt , and ratio between the time step and the computational time τ consumed to propagate the quantum-dynamics a single step for the methane molecule. Δt is given in femtoseconds and $\Delta t/\tau$ in femtoseconds/second.

Magnus (NC = 5)		Magnus (NC = 10)		Magnus (NC = 15)		Magnus (NC = 30)		Verlet	
Δt	$\Delta t/\tau$	Δt	$\Delta t/\tau$	Δt	$\Delta t/\tau$	Δt	$\Delta t/\tau$	Δt	$\Delta t/\tau$
0.002	0.013	0.007	0.044	0.010	0.063	0.012	0.076	0.001	0.013

density matrix, but through a significant decrease in the propagation time step.

In Table III the parameter τ represents the computational time to propagate the dynamics one time step, and therefore the ratio $\Delta t/\tau$ gives a measurement for the efficiency of the propagation algorithm. For a small molecule as methane, the Verlet algorithm turns out to be nearly twice as fast as the Magnus propagator, but the time step for the later is nearly 12 times larger. As a result, for this system the Magnus expansion excels the performance of the Verlet scheme by a factor of 6. The excess in computational GPU time for increasing the number of commutators in the first-order Magnus expansion is negligible.

C. Solvatochromic shift in formamide

Formamide represents the simplest model for the peptidic bond. The interest in the application of UV spectroscopy to the structural, dynamical, and electronic characterization of the peptidic linkage in proteins, has motivated a lot of experimental²⁷⁻³¹ and theoretical³²⁻⁴¹ research on the excitations and absorption of formamide, acetamide, and other related molecules. The UV spectrum of formamide is characterized by an intense band at 7.4 eV involving non-bonding and antibonding π orbitals (denoted $\pi_{nb} \rightarrow \pi^*$),^{27,30} which experiences a red-shift of 0.5 eV in aqueous solution.^{28,29} Other weaker signals are also present, attributed to $n \rightarrow \pi^*$ transitions and to Rydberg excitations. Various authors have applied electronic structure methodologies to interpret the origin of these bands, and to examine the effect of the solvent on their energies. Early simulations of the absorption spectra of small amides in the gas phase or hydrated by a few water molecules, were based on static calculations with the CASSCF and CASPT2 methods. The effect of finite temperature was modelled in water solution by several workers like Mennucci or Besley and Hirst, who performed TDDFT calculations embedded in a continuum solvent on an ensemble of configurations generated through classical molecular dynamics.³⁵⁻³⁷ Besley and Doltsinis followed this same idea to obtain the absorption spectra of formamide in the gas phase, combining in this case Car-Parrinello molecular dynamics with MRCI to extract the electronic transitions.³⁹ Very recently, Gordon *et al.* have interpreted the observed solvatochromic shifts in terms of the solvent effect on the HOMO and LUMO energies.⁴¹ In particular, they explored the finite temperature broadening of the spectral line shape both in vacuum and in solution, by performing molecular dynamics simulations with DFT and then computing the spectra at the TDDFT/PBE0 level for a set of configurations. The solvent environment was modelled in a QM-MM framework with the

inclusion of 100 H₂O molecules represented through the effective fragment potential EFP1.⁴¹ Based on this approach, these authors reported absorption energies of 8.74 eV and 8.31 eV in the gas and in the aqueous phases, respectively.

We have examined this problem applying our QM-MM methodology. More specifically, we have run molecular dynamics simulations of a formamide molecule in an octahedral simulation box of radius 15 Å containing 1067 water molecules in periodic boundary conditions, to analyze the displacement of the $\pi_{nb} \rightarrow \pi^*$ band in solution. The configurational sampling of solvated formamide was achieved through QM-MM Born-Oppenheimer molecular dynamics, which combined DFT and the TIP3P water force-field to describe the H₂O molecules.⁴² This scheme was applied to propagate the positions of the nuclei adiabatically on the Born-Oppenheimer surface, coupled to a Langevin thermostat. A set of configurations extracted from these QM-MM trajectories were then computed at the RT-TDDFT level, using exactly the same PBE/TIP3P hamiltonian as the one adopted to perform the molecular dynamics. Hence, an interesting feature of our approach is that it allows to use exactly the same QM-MM hamiltonian (combining PBE + Amber) for both the Born-Oppenheimer and the electron dynamics. The spectra of formamide were thus computed as an average of different configurations, each of which was extracted from a trajectory thermalized during at least 5 ps, preceded by an equilibration scheme including a previous QM-MM thermalization at the AM1 level. Figure 3 displays the averaged spectra in vacuum and in water. The position of the absorption maximum appears to converge for a set of less than 20 configurations in

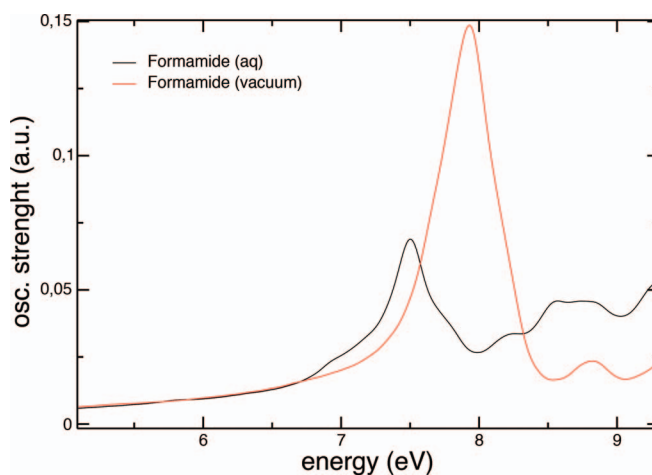


FIG. 3. RT-TDDFT simulated spectra of formamide in the gas phase and in aqueous solution. The results correspond to averages taken over 20 and 30 configurations, respectively, generated through molecular dynamics at 300 K. QM-MM simulations were performed to obtain the spectra in solution.

the gas phase, whereas in solution it requires a more extensive sampling, involving 30 different snapshots. This should be expected since the solvated formamide explores many more local minima than in the gas phase. We note that the starting geometries for the different structures employed in the construction of the TDDFT spectra, have their origin in a classical molecular dynamics with the Amber force-field extended over 15 ns. This ensures that in the frames taken to average the RT-TDDFT spectra, the solvent configurations are well decorrelated from one another. The calculated $\pi_{nb} \rightarrow \pi^*$ band exhibits a deviation of approximately +0.6 eV with respect to the experimental value, owing to the DFT description (errors in the energies predicted by TDDFT for the various excitations in the amide group were shown to be in this range, depending on the exchange-correlation functional³⁷). More importantly, our approach yields a solvent induced red-shift of 0.46 eV, remarkably reproducing the experimental value of 0.5 eV. This result proves that the present scheme is appropriate to study the UV spectroscopy of the peptidic bond in aqueous solution, suggesting at the same time that the participation of the solvent in the solvatochromic shift of the amide

is purely electrostatic. In other words, it does not seem to be necessary to have a quantum-mechanical description of the solvation shell to attain a quantitatively accurate representation of this phenomena.

D. The effect of the protein environment on the Soret band of heme

Porphyrins exhibit two kinds of well characterized electronic transitions in the visible region, associated with the so called Q and Soret bands.⁴³ The last ones have been extensively utilized to provide rich morphologic and biochemical information about the environment of the heme group such as the type and number of ligands coordinated to the metal atom or its oxidation state.⁴⁴⁻⁴⁶ In this section we apply our TDDFT QM-MM scheme to investigate the Soret band of heme in the Flavohemoglobin of Escherichia Coli (ecFlavoHb).⁴⁷ Flavohemoglobins (FlavoHbs) are a family of heme proteins widely distributed among bacteria and yeasts. Their structural features involve a globin domain containing a penta-coordinated heme group (with an axial imidazole

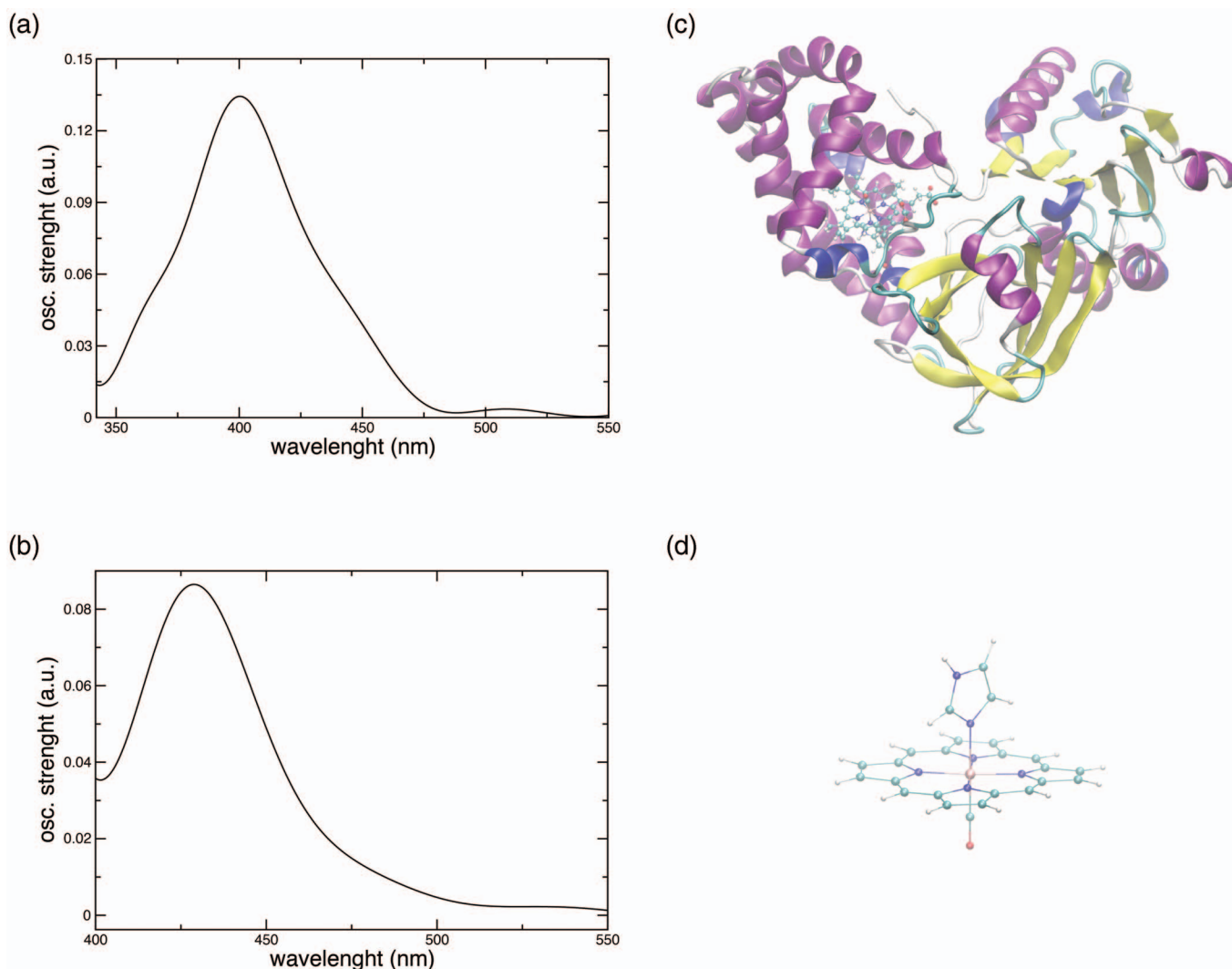


FIG. 4. Simulated spectra of hexa-coordinated heme in Flavohemoglobin (a) and in the gas phase (b). The corresponding model structures are represented in (c) and (d). The spectra in Flavohemoglobin was obtained using QM-MM calculations.

ligand), fused with a ferredoxin reductaslike module. Given their high redox reactivity, different physiological roles associated with cell response to oxidative and nitrosative stress have been proposed.⁴⁷ In particular, ecFlavoHb consists of 396 aminoacids plus a heme group and a flavin-adenine dinucleotide (FAD) molecule. A model structure for this system is displayed in the upper right panel of Figure 4. The total number of atoms is 6270, which makes a full quantum mechanics treatment unaffordable.

The absorption spectrum of CO bound ecFlavoHb was computed on a partially optimized structure, including in the QM region the heme group plus the two axial ligands: imidazole and carbon monoxide. The Soret band is depicted in the upper left panel of Figure 4. On the other hand, the absorption peak corresponding to the isolated hexa-coordinated heme group with the same ligands is shown in the lower panel. These bands are centered at 400.6 nm and 436.0 nm for the protein and the isolated complex, respectively. It is important to emphasize that in both the ecFlavoHb and the isolated heme models, the QM region is exactly the same. Therefore, the observed blue shift of ~ 35 nm in the Soret band results from the incorporation of the MM subsystem representing the surrounding protein.

The Soret band for the penta-coordinated heme in ecFlavoHb has been experimentally characterized at 403 nm.⁴⁷ This can be considered as an upper bound to the position of the Soret peak in CO ligated ecFlavoHb, since the coordination of CO to a penta-coordinated heme produces a small blue shift in the Soret absorption, typically in the range of 5–15 nm.⁴⁸ Then, despite the lack of a direct experimental comparison, it can be concluded that the observed shift in the ecFlavoHb heme with respect to the gas phase hexa-coordinated porphyrin is qualitatively correct. We note that these results must be considered only qualitative, as they proceed from static configurations of the heme model. As discussed in Ref. 49, the convergence of the spectrum with respect to the number of sampled configurations is mandatory if an accurate prediction is sought.

IV. CONCLUSIONS

We have developed a powerful scheme to perform real-time TDDFT electron dynamics in a QM-MM framework. This implementation can easily handle systems in the order of 100 quantum atoms surrounded by thousands of classical nuclei, enabling the investigation of the effect of a complex environment, such as a solvent or a protein matrix, on the spectroscopy, the conductance, or the out-of-equilibrium behavior of large molecular species. The most demanding parts of the present scheme involve the exchange-correlation energy, which is computed in GPU and scales linearly with the number of atoms,¹⁷ and the $[\mathbf{H}, \mathbf{P}]$ commutator. The later is also efficiently handled in GPU, hence representing a negligible fraction of the total computing time, providing the dimension of the matrices is not huge. As a consequence, the global scaling turns out to be almost linear. Since the most common LR-TDDFT implementations scale with N^3 or N^4 (with N the number of electrons), the parallelization in GPU renders this method competitive with respect to the linear re-

sponse approach to calculate the electronic transitions of large molecular systems, in the order of 100 atoms or more.

The Magnus propagator was found to be more efficient than the Verlet algorithm for small systems by a factor of 6. The inclusion of additional commutators in the BCH scheme, up to $NC \approx 30$, allows for a longer time step at a negligible GPU computation cost. The presence of heavy atoms severely restrains the maximum time step, because of the high frequencies associated with the inner electronic states.

In general, the observed quantum response of the system arises from an average on an ensemble of instantaneous configurations, which can be obtained, for example, from a molecular dynamics trajectory. An interesting feature of the present method is that it allows to use exactly the same hamiltonian to sample the configurational space through molecular dynamics simulations, and to propagate the time-dependent equations for the electronic states. This procedure has been illustrated for the case of formamide in aqueous solution, yielding results in excellent agreement with experiments. Beyond the calculation of absorption spectra, the code reported here may be useful to study charge transport in biological environments. For example, the electron transfer rates between enzymatic active sites are largely determined by the geometries adopted by the peptidic chains in between, and modulated by the surrounding residues. A reliable estimation of the transport properties in this complex would require an extensive averaging on the accessible structures of the system, attainable via QM-MM molecular dynamics simulations. The methodology described in this article is particularly well suited to address this kind of problems, which will be the subject of future work.

ACKNOWLEDGMENTS

We express our gratitude to Professor Dario Estrin for valuable discussions. We are also grateful to the reviewers, for their thorough reading of our work and for very helpful comments. This study has been in part supported by grants of the Agencia Nacional de Promocion Cientifica y Tecnologica de Argentina, PICT 2007-2111 and PICT 2012-2292, and UBACYT 20020120100333BA. U.M. acknowledges CONICET for a doctoral fellowship.

¹M. A. L. Marques, C. A. Ullrich, F. Nogueira, A. Rubio, K. Burke, and E. K. U. Gross, *Time-Dependent Density Functional Theory* (Springer, 2006).

²M. Casida and M. Huix-Rotllant, *Annu. Rev. Phys. Chem.* **63**, 287 (2012).

³M. Petersilka, U. J. Gossmann, and E. K. U. Gross, *Phys. Rev. Lett.* **76**, 1212 (1996).

⁴M. E. Casida, *Recent Advances in Density-Functional Methods* (World Scientific, 1995).

⁵C. Jamorski, M. E. Casida, and D. R. Salahub, *J. Chem. Phys.* **104**, 5134 (1996).

⁶K. Yabana and G. F. Bertsch, *Phys. Rev. B* **54**, 4484 (1996).

⁷A. Castro, H. Appel, M. Oliveira, C. Rozzi, X. Andrade, F. Lorenzen, M. Marques, E. Gross, and A. Rubio, *Phys. Status Solidi B* **243**, 2465 (2006).

⁸O. Sugino and Y. Miyamoto, *Phys. Rev. B* **59**, 2579 (1999).

⁹X. Qian, J. Li, X. Lin, and S. Yip, *Phys. Rev. B* **73**, 035408 (2006).

¹⁰A. Tsolakidis, D. Sánchez-Portal, and R. M. Martin, *Phys. Rev. B* **66**, 235416 (2002).

¹¹C.-L. Cheng, J. S. Evans, and T. Van Voorhis, *Phys. Rev. B* **74**, 155112 (2006).

¹²X. Li and J. C. Tully, *Chem. Phys. Lett.* **439**, 199 (2007).

- ¹³K. Lopata and N. Govind, *J. Chem. Theory Comput.* **7**, 1344 (2011).
- ¹⁴F. Wang, C. Y. Yam, G. Chen, X. Wang, K. Fan, T. A. Niehaus, and T. Frauenheim, *Phys. Rev. B* **76**, 045114 (2007).
- ¹⁵M. B. Oviedo, C. F. A. Negre, and C. G. Sanchez, *Phys. Chem. Chem. Phys.* **12**, 6706 (2010).
- ¹⁶J. Jakowski and K. Morokuma, *J. Chem. Phys.* **130**, 224106 (2009).
- ¹⁷M. A. Nitsche, M. Ferreria, E. E. Mocskos, and M. C. Gonzalez-Lebrero, *J. Chem. Theory Comput.* **10**, 959 (2014).
- ¹⁸D. A. Case *et al.*, *Amber 11*, University of California, San Francisco, 2011; see <http://ambermd.org/>.
- ¹⁹D. J. Tannor, *Introduction to Quantum Mechanics. A Time-Dependent Perspective* (University Science Books, 2007).
- ²⁰W. Magnus, *Commun. Pure Appl. Math.* **7**, 649 (1954).
- ²¹H. Chen, J. M. McMahon, M. A. Ratner, and G. C. Schatz, *J. Phys. Chem. C* **114**, 14384 (2010).
- ²²D. J. Masiello and G. C. Schatz, *J. Chem. Phys.* **132**, 064102 (2010).
- ²³J. P. Perdew, K. Burke, and M. Ernzerhof, *Phys. Rev. Lett.* **77**, 3865 (1996).
- ²⁴N. Godbout, D. R. Salahub, J. Andzelm, and E. Wimmer, *Can. J. Chem.* **70**, 560 (1992).
- ²⁵F. Neese *et al.*, *Orca 2.9*, Max-Planck-Institute for Inorganic Chemistry, 2012. see <http://cec.mpg.de/forum/>.
- ²⁶M. J. Frisch, G. W. Trucks, H. B. Schlegel *et al.*, Gaussian 03, Revision C.02, Gaussian, Inc., Wallingford, CT, 2004.
- ²⁷H. D. Hunt and W. T. Simpson, *J. Am. Chem. Soc.* **75**, 4540 (1953).
- ²⁸E. B. Nielsen and J. A. Schellman, *J. Phys. Chem.* **71**, 2297 (1967).
- ²⁹H. Basch, M. B. Robin, and N. A. Kuebler, *J. Chem. Phys.* **49**, 5007 (1968).
- ³⁰J. M. Gingell, N. J. Mason, H. Zhao, I. C. Walker, and M. R. Siggel, *Chem. Phys.* **220**, 191 (1997).
- ³¹T. Goto, A. Ikehata, Y. Morisawa, and Y. Ozaki, *J. Phys. Chem. A* **117**, 2517 (2013).
- ³²L. Serrano-Andrés and M. P. Fülscher, *J. Am. Chem. Soc.* **118**, 12190 (1996).
- ³³N. A. Besley and J. D. Hirst, *J. Phys. Chem. A* **102**, 10791 (1998).
- ³⁴N. A. Besley and J. D. Hirst, *J. Am. Chem. Soc.* **121**, 8559 (1999).
- ³⁵N. A. Besley, M. T. Oakley, A. J. Cowan, and J. D. Hirst, *J. Am. Chem. Soc.* **126**, 13502 (2004).
- ³⁶B. Mennucci and J. M. Martínez, *J. Phys. Chem. B* **109**, 9818 (2005).
- ³⁷J. Šebek, Z. Kejík, and P. Bouř, *J. Phys. Chem. A* **110**, 4702 (2006).
- ³⁸A. Chen, X. Pu, S. He, Y. Guo, Z. Wen, M. Li, N.-B. Wong, and A. Tian, *New J. Chem.* **33**, 1709 (2009).
- ³⁹N. A. Besley and N. L. Doltsinis, *J. Chem. Theory Comput.* **2**, 1598 (2006).
- ⁴⁰J. Jiang, D. Abramavicius, B. M. Bulheller, J. D. Hirst, and S. Mukamel, *J. Phys. Chem. B* **114**, 8270 (2010).
- ⁴¹N. De Silva, S. Y. Willow, and M. S. Gordon, *J. Phys. Chem. A* **117**, 11847 (2013).
- ⁴²M. W. Mahoney and W. L. Jorgensen, *J. Chem. Phys.* **112**, 8910 (2000).
- ⁴³*The Porphyrins*, edited by D. Dolphin (Academic Press, 1978), Vol. III.
- ⁴⁴C. Rimington, *Biochem. J.* **75**, 620 (1960).
- ⁴⁵M. Momenteau and C. A. Reed, *Chem. Rev.* **94**, 659 (1994).
- ⁴⁶G. H. Loew and D. L. Harris, *Chem. Rev.* **100**, 407 (2000).
- ⁴⁷A. Ilari, A. Bonamore, A. Farina, K. A. Johnson, and A. Boffi, *J. Biol. Chem.* **277**, 23725 (2002).
- ⁴⁸W. G. Eisert, E. O. Degenkolb, L. J. Noe, and P. M. Rentzepis, *Biophys. J.* **25**, 455 (1979).
- ⁴⁹O. Valsson, P. Campomanes, I. Tavernelli, U. Rothlisberger, and C. Filippi, *J. Chem. Theory Comput.* **9**, 2441 (2013).