

Quantum Calculations in Solution for Large to Very Large Molecules: A New Linear Scaling QM/Continuum Approach

Filippo Lipparini,^{*,†,‡,§,¶} Louis Lagardère,[¶] Giovanni Scalmani,[§] Benjamin Stamm,^{†,||} Eric Cancès,[⊥] Yvon Maday,^{†,‡,∇} Jean-Philip Piquemal,^{†,||} Michael J. Frisch,[§] and Benedetta Menucci[◆]

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

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

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

[§]Gaussian, Inc., 340 Quinipiac Street, Building 40, Wallingford, Connecticut 06492, United States

^{||}CNRS, UMR 7598 and 7616, F-75005 Paris, France

[⊥]CERMICS, Ecole des Ponts and INRIA, Université Paris-Est, 6 & 8 avenue Blaise Pascal, 77455 Marne-la-Vallée Cedex 2, France

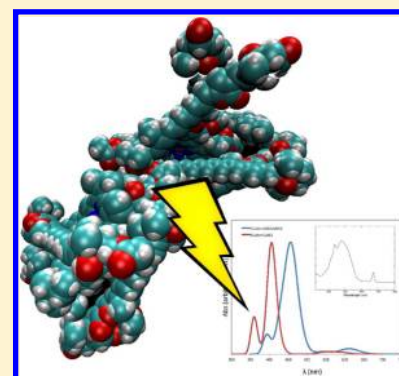
[#]Institut Universitaire de France, France

[∇]Division of Applied Maths, Brown University, Providence, Rhode Island 02912, United States

[◆]Dipartimento di Chimica e Chimica Industriale, Università di Pisa, Via Risorgimento 35, 56126 Pisa, Italy

S Supporting Information

ABSTRACT: We present a new implementation of continuum solvation models for semiempirical Hamiltonians that allows the description of environmental effects on very large molecular systems. In this approach based on a domain decomposition strategy of the COSMO model (ddCOSMO), the solution to the COSMO equations is no longer the computational bottleneck but becomes a negligible part of the overall computation time. In this Letter, we analyze the computational impact of COSMO on the solution of the SCF equations for large to very large molecules, using semiempirical Hamiltonians, for both the new ddCOSMO implementation and the most recent, linear scaling one, based on the fast multipole method. A further analysis is on the simulation of the UV/visible spectrum of a light-harvesting pigment–protein complex. All of the results show how the new ddCOSMO algorithm paves the way to routine computations for large molecular systems in the condensed phase.



SECTION: Molecular Structure, Quantum Chemistry, and General Theory

In the last decades, quantum mechanical (QM) simulations of molecular properties and processes have moved from the realm of isolated (gas-phase) molecules to that of by far more interesting solvated systems. This progress has been largely made possible due to the coupling of QM approaches to classical solvation models. Such a coupling has followed two main strategies; either the classical environment has been described through molecular mechanics (MM) force fields, or it has been modeled as a continuum dielectric.^{1–10} In particular, the second strategy presents many advantages with respect to the first one, especially when they are both coupled to QM descriptions. Continuum models in fact do not require the knowledge of a suited force field to describe the solute–solvent interaction but only to define the boundary between solute and solvent; furthermore, they implicitly include statistical averages, which instead have to be explicitly taken into account in terms of many different configurations of the solvent when using QM/MM formulations. Both of these two aspects have made continuum solvation models very popular, and nowadays, almost all of the main software packages have at least one

continuum solvation model as a possible option. The rapid expansion of QM/continuum approaches has however shown some intrinsic limitations both from the physical and the numerical point of view. Physically, continuum models cannot properly account for specific solute–solvent interactions, among which hydrogen bonding is the most common and important example. However, this limit can be rather easily solved by extending the QM part of the system so as to include some solvent molecules or using a mixed QM/MM/continuum formulation.^{11–17}

The numerical limitations are instead more difficult to be solved even if different effective strategies have been proposed so far. To understand these numerical issues, we have to recall that the implementation of continuum solvation models relies on the definition of a “cavity” that embeds the QM part of the system and on the introduction of a numerical procedure to

Received: February 5, 2014

Accepted: February 25, 2014

solve the resulting three-dimensional partial differential (Poisson's) equations. Such equations can be solved with different approaches. The most popular versions of QM/continuum models use either the so-called dielectric or conductor boundary conditions (or jumps) at the cavity surface. The two alternative strategies have led to the two main families of modern continuum models known as the PCM (polarizable continuum model) and COSMO (conductor-like screening model) (see two recent reviews for a more complete list of references on the two approaches).^{18,19} In all cases, the differential equations are recast as integral equations on the cavity boundary, which can be treated with standard numerical techniques such as the boundary element method (BEM) through which the cavity surface is discretized using a mesh.

The BEM is a very effective approach, but it suffers from two major disadvantages. First, as the integral operator is nonlocal (it corresponds to the kernel of the Laplacian), the matrix resulting from the discretization of the integral equation is dense, and all of its N^2 elements are, in principle, nonzero. This implies $O(N^2)$ memory requirements and $O(N^3)$ or $O(N^2)$ computational cost to solve the linear system by matrix inversion or with an iterative procedure, where no sparsity can be exploited when computing the necessary matrix/vector products. For this last case, it is also to be noticed that if enough memory is available, the matrix can be built and stored; if this is not the case (on standard machines, the limit is roughly 1000 atoms), the matrix has to be assembled on-the-fly at each matrix vector product, further increasing the computational cost. Furthermore, The BEM matrix is usually very ill-conditioned; iterative schemes struggle to achieve convergence. Second, as moving the atoms causes the mesh of the external surface to change and surface elements to become buried in the inside of the cavity or exposed, the energy is not smooth with respect to the nuclei positions, making the use of continuum solvation models for geometry optimizations or molecular dynamics simulations either hard or impossible.

In the past decade, a lot of progress has been done on both shortcomings. As the matrix/vector products necessary for any iterative procedure involve the computation of either a potential or an electric field, fast summation techniques like the fast multipole method²⁰ (FMM) can be successfully employed in order to achieve linear scaling of both computational cost and memory requirements; the FMM was implemented in the context of continuum solvation by Scalmani and co-workers in 2004.²¹ Furthermore, a different strategy to solve the PCM/COSMO equations based on a variational approach^{22,23} was presented by some of us with interesting computational results. To overcome the continuity problem, more advanced discretization techniques can be used in order to obtain a smooth solvation energy as a function of the atomic coordinates; a discretization scheme based on Lebedev grids, Gaussian basis functions, and switching functions was proposed for COSMO by York and Karplus in 1999²⁴ and extended and generalized independently by Scalmani and Frisch²⁵ and by Lange and Herbert^{26,27} in 2010. By using the so-called continuous FMM to compute the Coulomb interaction between Gaussian densities,²⁸ it is possible to achieve linear scaling in computational cost also for continuous implementations.

Recently, a completely different strategy to solve the COSMO problem has been presented by some of us.^{29,30} Such a new discretization of the COSMO equations, which we will refer to as ddCOSMO, is based on Schwarz's domain

decomposition method and has been proven to be both smooth and fast; furthermore, linear scaling in both computational cost and memory requirements with respect to the system's size is implicit in the procedure without needing to resort to fast summation techniques. In two previous papers, the discretization has been thoroughly analyzed from a mathematical point of view,²⁹ and an efficient and parallel implementation has been proposed³⁰ for a solute described at a classical level of theory. In this Letter, we present the results obtained by the first coupling of the ddCOSMO method with QM calculations.

As the focus is here on very large systems, the QM approach is based on semiempirical Hamiltonians.^{31–35} In the framework of semiempirical methods, the electrostatic properties of the molecular system are modeled with a collection of atomic point multipoles $[\Theta_j]_l^m$ (usually up to the quadrupole term), which are related to the electronic density matrix by means of a parametrized linear transformation³⁶

$$[\Theta_j]_l^m = \sum_{\mu\nu} \Lambda_{jlm,\mu\nu} P_{\mu\nu} \quad (1)$$

where l and m are multipolar indexes, j refers to an atom, and the sum over $\mu\nu$ runs over the atomic orbitals (AOs). The ddCOSMO solvation energy assumes, in this particular case, a very simple expression

$$E^s(\Theta) = \frac{1}{2} f(\epsilon) \sum_{j=1}^M \sum_{l=0}^2 \sum_{m=-l}^l \sqrt{\frac{4\pi}{2l+1}} [\Theta_j]_l^m [\sigma_j]_l^m \quad (2)$$

Here, $f(\epsilon)$ is a constant scaling function used to account for the nonconductor nature of a solvent, M is the number of atoms, and $[\sigma_j]_l^m$ is the l,m coefficient of the solution to the ddCOSMO linear equations $L\sigma = g$ on the j th sphere, where the right-hand side $[g_j]_l^m$ is the molecular electrostatic potential weighted with the proper switching factors. All of the details on the ddCOSMO algorithm can be found in refs 29 and 30, and more details are also included in the Supporting Information. In order to couple ddCOSMO to a semiempirical calculation, a Fock matrix contribution needs to be computed. Such a contribution can be computed by differentiating the energy with respect to the atomic multipoles and then using eq 1 to transform the atomic Fock matrix in the AO basis

$$\begin{aligned} [\Xi_i]_l^m &= \frac{\partial E^s}{\partial [\Theta_i]_l^m} \\ &= \frac{1}{2} f(\epsilon) \left[\sqrt{\frac{4\pi}{2l+1}} [\sigma_i]_l^m \delta_{l \leq 2} + \sum_{j=1}^M \sum_{l'm'} [s_j]_{l'm'}^m \frac{\partial [g_j]_{l'm'}^m}{\partial [\Theta_i]_l^m} \right] \end{aligned} \quad (3)$$

where $[s_j]_{l'm'}^m$ is the l,m coefficient of the solution to the ddCOSMO adjoint problem on the j th sphere and $\delta_{l \leq 2} = 1$ if $0 \leq l \leq 2$ and vanishes otherwise. Finally

$$F_{\mu\nu} = \sum_{i=1}^M \sum_{lm} \Lambda_{\mu\nu,ilm}^\dagger [\Xi_i]_l^m \quad (4)$$

A complete formulation of ddCOSMO for general QM solutes will be the object of a future communication.

To summarize, the ingredients of a semiempirical QM/ddCOSMO implementation are the following. To assemble the right-hand side g of the ddCOSMO linear system, one needs a parametrized linear transformation³⁶ between the density matrix and the set of atomic multipoles (see eq 1), which is defined by the specific semiempirical method used and contains

the semiempirical parameters, the solute's potential at the cavity, which is computed as the potential produced by the aforementioned atomic multipoles, and some geometrical quantities that are related to the ddCOSMO regularization. Then, to compute the energy given a density, one has to solve the ddCOSMO linear system and use eq 2; finally, to assemble the ddCOSMO Fock matrix contribution, the adjoint ddCOSMO linear system has to be solved, and the two contributions as in eq 3 have to be assembled (see the Supporting Information for technical details). The procedure here described applies with slight differences also to the ddCOSMO contributions to response theory; as a last remark, in order to compute the forces, either the derivatives of the solute's potential, which are standard quantities, or the derivatives of the ddCOSMO matrix need to be computed; these latter have been thoroughly described elsewhere.³⁰

The standard COSMO (or C-PCM) implementation is somewhat simplified by the fact that the matrix arising from the discretization of the COSMO integral equation is symmetric; as a consequence, there is no need to compute the solution to the adjoint linear system, and the Fock matrix contribution can be assembled by contracting the solution to the discretized equation with the potential integrals. In the following, we will compare the computational effort required by both the ddCOSMO and the continuum surface charge²⁵ (CSC) discretization in order to solve the discretized equations and assemble the Fock matrix contribution, which for ddCOSMO only includes the effort of solving the adjoint system. Before doing so, two remarks are mandatory. First, while for ddCOSMO assembling the potential is straightforward, the CSC implementation discretizes the apparent surface charge (ASC) in terms of spherical Gaussian functions at the cavity, and thus, to compute the potential and the Fock matrix contributions, integrals such as

$$\Phi_i^n = \int_{\mathbb{R}^3} dx \int_{\mathbb{R}^3} dy \frac{\rho(\mathbf{x})\phi(\mathbf{y} - \mathbf{s}_i^n)}{|\mathbf{x} - \mathbf{y}|} \quad (5)$$

where $\phi(\cdot - \mathbf{s}_i^n)$ is a Gaussian function centered at the point \mathbf{s}_i^n on the cavity, have to be evaluated. In order to use the standard machinery to compute bielectronic integrals and hence exploit its efficiency, it is possible to approximate the point-multipolar distribution $\rho(\mathbf{x})$ with a collection of very spiked (i.e., with exponents larger than 10^8) Gaussian basis functions of different angular momentum (i.e., s, p, and d for standard semiempirical methods). Notice that the linear scaling techniques used for Coulomb integrals, such as the continuous FMM, can be successfully exploited. Second, the use of the FMM to compute the matrix–vector products for the iterative solution of the CSC equations introduces a further approximation with respect to the ones implied by the discretization and two additional parameters, that is, the maximum angular momentum in the multipolar expansion and the size of the smallest box. Furthermore, the performances of the FMM depend on the geometry of the system, with the best performances for long, linear molecules and the worst for globular one; this is not the case for ddCOSMO, where the performances depend more on the molecular topology, that is, on how many neighbors each atom has; such a number (between 15 and 25 for van der Waals cavities) is not affected by the overall molecular shape.

The computational cost of solving the linear system arising from the COSMO discretization depends on the size of the matrix, its characteristics, and the numerical procedure, that is,

matrix inversion or iterative solution. This step discriminates the various algorithms as the size of the matrix is given by the number of surface elements (N_{Ts}) for CSC and by the number of basis functions (N_{b}) for ddCOSMO, which is given by the number of atoms times the number of spherical harmonics used to expand the local intermediates. Through all of the simulations, we use $L_{\text{max}} = 10$, that is, 121 spherical harmonics per sphere. These quantities are reported in Table 1 for some medium- to large-sized systems and can be used to estimate the overall cost of the computation.

Table 1. Size of the Systems Used As Benchmark^a

system	M	N_{g}	N_{b}
vancomycin	377	22294	45617
Hiv-1-GP41	530	28931	64130
l-plectasin	567	31154	68607
glutaredoxin	1277	66996	154517
UBCHSB	2360	100937	285560
carboxylase	6605	272205	799205

^aNumber of atoms (M); for CSC, the number of surface elements (N_{g}); for ddCOSMO, the number of basis functions (N_{b}).

We report in Table 2 the timings of both the solution to the linear system (labeled σ) and the assembling of the Fock matrix contribution (labeled F) for the various algorithms and for the chosen systems. A scaled van der Waals cavity was used for all of the system, and the solutes were described with the AM1 Hamiltonian.³¹ For completeness, we also report the timings for the standard CSC implementation, that is, where the linear equations are solved by matrix inversion, if feasible, or with an iterative method without exploiting the FMM machinery. All of the calculations have been performed on a cluster node equipped with 2 Xeon E5-2650 processors with 8 cores each operating at 2.0 GHz and 64 GB of RAM operating at 1.6 GHz. A locally modified development version of the Gaussian suite of programs³⁷ has been used for all of the simulations. The efficiency of the new ddCOSMO procedure is clearly demonstrated by the reported timings. Of course, as it scales linearly within computational cost with respect to the size of the system, ddCOSMO is, for large enough systems faster than CSC if the $O(N^3)$ matrix inversion or the $O(N^2)$ iterative procedure is used; the interesting comparison is with the linear scaling CSC/FMM implementation. While the CSC/FMM is already very fast if compared with the standard iterative procedure, the time needed to solve just once the linear equations of the solvation model using CSC/FMM compares with the time for a SCF step for the largest system or even with the overall computation in vacuo for the other ones; the total computational effort to solve the SCF problem for the solvated system is by far larger than the time needed to perform the same computation for the isolated molecule. It is also interesting to notice that the crossover between the FMM implementation and the standard, quadratic one is reached only for considerably large molecules; for vancomycin (377 atoms), the $O(N^2)$ algorithm is faster than the linear scaling one (20 versus 43 s).

The same is not true for ddCOSMO, which is 1–2 orders of magnitude faster than the CSC/FMM implementation. This is the result of two factors; the ddCOSMO matrix is block-sparse and better conditioned. The ddCOSMO matrix is made by $M \times M$ blocks of size $(L_{\text{max}} + 1)^2$, where we remind that we have

Table 2. Timings for the Solution of the C-PCM/COSMO Linear Equations and for the Computation of the Fock Operator Contribution for the Different Algorithms^a

system	CSC - inversion		CSC - iterative		CSC/FMM		ddCOSMO	
	σ	F	σ	F	σ	F	σ	F
vancomycin	15 min, 56 s	1 s	20 s	1 s	43 s	1 s	1 s	1 s
Hiv-1-GP41	34 min, 19 s	1 s	1 min, 26 s	1 s	57 s	1 s	1 s	1 s
l-pectasin	39 min, 21 s	1 s	1 min, 35 s	1 s	1 min, 17 s	1 s	1 s	1 s
glutaredoxin			8 min, 54 s	1 s	3 min, 2 s	1 s	2 s	3 s
glutaredoxin ^b			28 min, 43 s					
UBCH5B			94 min	4 s	6 min, 18 s	2 s	4 s	8 s
carboxylase			777 min	22 s	24 min, 42 s	3 s	9 s	13 s

^aAll the timings smaller than 1 s have been approximated to 1 s. ^bThe computation was repeated without the N^2 storage for CSC.

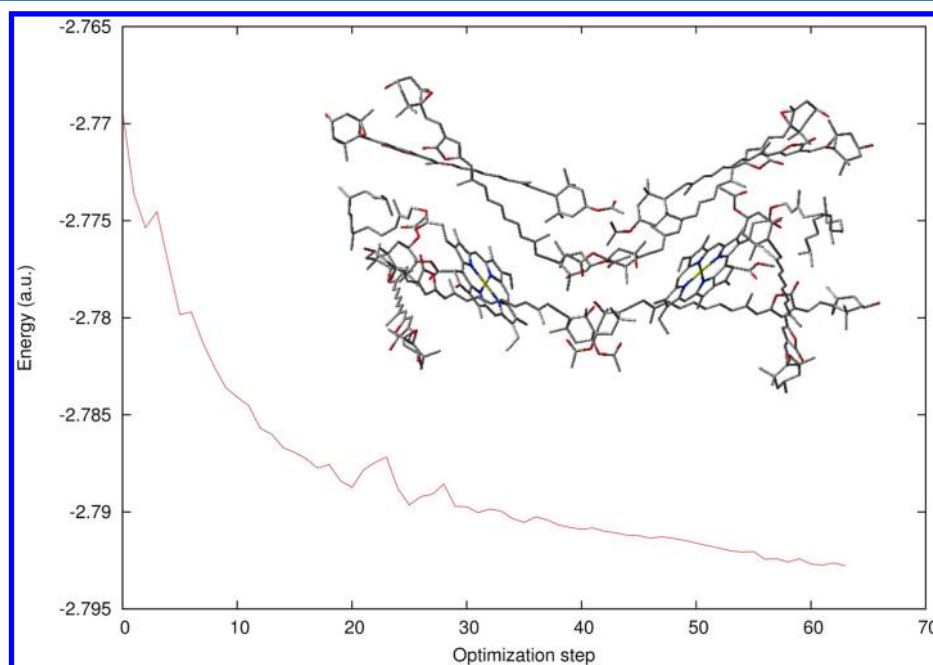


Figure 1. Energy profile of the geometry optimization of the PCP light-harvesting complex. The inset reports the optimized structure of the complex.

chosen $L_{\max} = 10$, and an off-diagonal block of the ddCOSMO matrix is nonzero only if it is associated with two intersecting spheres; for a large molecule, the large majority of the blocks are therefore zero, which makes the single iteration very fast. Furthermore, while the CSC matrix is ill-conditioned and a large number of iterations (usually, roughly 60–70) is required to get convergence, the ddCOSMO matrix is well-conditioned, and only 20–25 iterations are needed to achieve convergence.

In order to show the potentialities of the new algorithm, we have computed the UV/visible absorption spectrum of the light-harvesting complex of the peridinin chlorophyll protein (PCP), a water-soluble peripheral LH antenna from dinoflagellate *Amphidinium carterae*. PCP presents a trimeric quaternary structure with each monomer consisting of two domains connected by a loop and related by a two-fold pseudosymmetry axis.³⁸ Each domain encloses a central chlorophyll a (CHL) surrounded by four peridinin (PIDs) so that in each monomeric form, there are two CHLs and the eight PID pigments closely packed (see Figure 1). The measured absorption spectra consist of an intense absorption at 450–550 nm and a small band at 670 nm. The short-wavelength region is due to the PIDs and to the chlorophyll Soret bands, which overlap at the short-wavelength side,

whereas the long-wavelength region is due to the Qy chlorophyll band.

In the present analysis, we have simulated the entire complex (two CHLs and eight PIDs) at the QM level (namely, 1038 atoms), while the protein/solvent environment has been represented as an effective dielectric using COSMO with the static dielectric permittivity equal to $\epsilon = 15$ and the optical dielectric permittivity equal to $\epsilon_{\infty} = 2$.³⁹

We have first performed a geometry optimization of the complex starting from the crystal structure³⁸ using analytical gradients in combination with the AM1 Hamiltonian; the resulting geometry was used for simulating the UV/visible absorption spectrum at the ZINDO level. To avoid dealing with a massive number of intruder states, we have restricted the number of orbitals used to compute the excitation properties to the 20 highest occupied and 20 lowest virtual. A non-equilibrium solvation regime was assumed for the vertical excitation.

The energy profile of the geometry optimization is reported in Figure 1; the convergence was smooth and without trouble, and the computation could be carried out on a standard cluster node in roughly 9 h. Notice that, using the ratio that one can

deduct from Table 2, this same computation would take roughly 1 month if performed using the FMM machinery.

The excited-state computation only took a few minutes, and the spectrum obtained by convoluting gaussians of 0.15 eV half-width at half-height centered at the excitation frequencies computed with ZINDO is reported in Figure 2. The

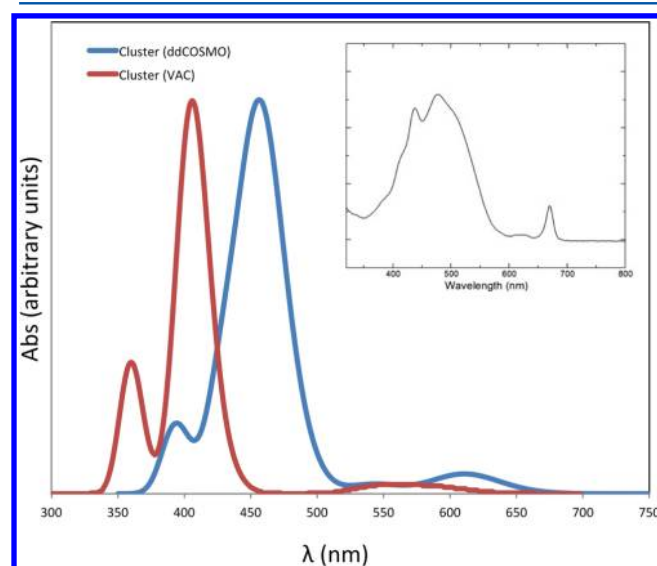


Figure 2. UV/vis spectrum of the PCP light-harvesting complex as computed with the ZINDO method in vacuo and with ddCOSMO. The spectra have been normalized with respect to the maxima. The inset shows the experimental spectrum extracted from ref 40.

comparison shows a qualitatively correct agreement; all of the observed bands are reproduced even if the whole calculated spectrum is blue-shifted with respect to the experimental one. Indeed, to accurately reproduce the experimental spectrum, a better QM level as well as the introduction of vibronic effects would be necessary. However, it is worth noting that the inclusion of the environmental effects through the ddCOSMO red shifts the spectrum with respect to that calculated in vacuo. This shows the importance that environment effects have in the simulation of the spectroscopic properties of light-harvesting complexes.

In this Letter, we have reported on a new algorithm to solve the COSMO equations that is characterized by several promising features. The algorithm is robust, numerically stable, and under variational control; furthermore, the number of parameters needed to define the discretization is very limited; in particular, the only parameters are the size of the spherical harmonics basis, the number of grid points used for numerical integration, the convergence threshold for the iterative solver, and a regularization parameter. The scaling properties of ddCOSMO are implicit in the model and do not depend on the molecular shape; also, approximate fast summation techniques are not needed, which further reduces the number of approximations and parameters involved. However, the most promising feature of ddCOSMO is its computational efficiency; even when compared with an already fast implementation based on the FMM, its performances are remarkable as the total time required per each SCF cycle is reduced, even for very large systems, from several minutes to a few seconds. This feature is particularly important if the computation has to be repeated several times, for instance, in order to compute a statistical average for a large, flexible molecule or in a (ab initio)

molecular dynamics simulation; the latter, in particular, becomes now feasible for medium-to-large systems using semiempirical Hamiltonians and continuum solvation. In conclusion, ddCOSMO paves the way for the use of polarizable continuum models for large and very large systems, and its extension to analytical derivatives allows its application in many research fields, going from (bio)chemical reactivity, to spectroscopic properties of supermolecular systems, to electronic processes involving many interacting molecules such as energy and charge transfers in photosynthetic systems.

■ ASSOCIATED CONTENT

Supporting Information

Derivation of the ddCOSMO contributions to the Fock matrix for a semiempirical Hamiltonian. This material is available free of charge via the Internet at <http://pubs.acs.org>.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: filippo.lipparini@courriel.upmc.fr.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported in part by the ANR Manif and the French state funds managed by CALSIMLAB and the ANR within the Investissements d'Avenir program under Reference ANR-11-IDEX-0004-02.

■ REFERENCES

- (1) Cramer, C.; Truhlar, D. Implicit Solvation Models: Equilibria, Structure, Spectra, and Dynamics. *Chem. Rev.* **1999**, *99*, 2161–2200.
- (2) Tomasi, J.; Mennucci, B.; Cammi, R. Quantum Mechanical Continuum Solvation Models. *Chem. Rev.* **2005**, *105*, 2999–3093.
- (3) Friesner, R. A.; Guallar, V. Ab Initio Quantum Chemical and Mixed Quantum Mechanics/Molecular Mechanics (QM/MM) Methods for Studying Enzymatic Catalysis. *Annu. Rev. Phys. Chem.* **2005**, *56*, 389–427.
- (4) Lin, H.; Truhlar, D. G. QM/MM: What Have We Learned, Where Are We, and Where Do We Go from Here? *Theor. Chem. Acc.* **2007**, *117*, 185–199.
- (5) Senn, H. M.; Thiel, W. QM/MM Methods for Biomolecular Systems. *Angew. Chem., Int. Ed.* **2009**, *48*, 1198–1229.
- (6) Hu, H.; Yang, W. Development and Application of Ab Initio QM/MM Methods for Mechanistic Simulation of Reactions in Solution and in Enzymes. *J. Mol. Struct.: THEOCHEM* **2009**, *898*, 17–30.
- (7) Barone, V.; Biczysko, M.; Brancato, G. In *Combining Quantum Mechanics and Molecular Mechanics. Some Recent Progresses in QM/MM Methods*; Sabin, J. R., Brandas, E., Eds.; Advances in Quantum Chemistry; Academic Press: New York, 2010; Vol. 59, pp 17–57.
- (8) Ferrer, S.; Ruiz-Pernia, J.; Marti, S.; Moliner, V.; Tunon, I.; Bertran, J.; Andres, J. In *Computational Chemistry Methods in Structural Biology*; Christov, C., Ed.; Advances in Protein Chemistry and Structural Biology; Academic Press: New York, 2011; Vol. 85, pp 81–142.
- (9) Olsen, J. M. H.; Kongsted, J. In *Molecular Properties through Polarizable Embedding*; Sabin, J. R., Brandas, E., Eds.; Advances in Quantum Chemistry; Academic Press: New York, 2011; Vol. 61, pp 107–143.
- (10) Mennucci, B. Modeling Environment Effects on Spectroscopies through QM/Classical Models. *Phys. Chem. Chem. Phys.* **2013**, *15*, 6583–6594.

- (11) Steindal, A. H.; Ruud, K.; Frediani, L.; Aidas, K.; Kongsted, J. Excitation Energies in Solution: The Fully Polarizable QM/MM/PCM Method. *J. Phys. Chem. B* **2011**, *115*, 3027–3037.
- (12) Lipparini, F.; Cappelli, C.; Barone, V. Linear Response Theory and Electronic Transition Energies for a Fully Polarizable QM/Classical Hamiltonian. *J. Chem. Theory Comput.* **2012**, *8*, 4153–4165.
- (13) Curutchet, C.; Muñoz-Losa, A.; Monti, S.; Kongsted, J.; Scholes, G. D.; Mennucci, B. Electronic Energy Transfer in Condensed Phase Studied by a Polarizable QM/MM Model. *J. Chem. Theory Comput.* **2009**, *5*, 1838–1848.
- (14) Caprasecca, S.; Curutchet, C.; Mennucci, B. Toward a Unified Modeling of Environment and Bridge-Mediated Contributions to Electronic Energy Transfer: A Fully Polarizable QM/MM/PCM Approach. *J. Chem. Theory Comput.* **2012**, *8*, 4462–4473.
- (15) Boulanger, E.; Thiel, W. Solvent Boundary Potentials for Hybrid QM/MM Computations Using Classical Drude Oscillators: A Fully Polarizable Model. *J. Chem. Theory Comput.* **2012**, *8*, 4527–4538.
- (16) Lipparini, F.; Cappelli, C.; Scalmani, G.; De Mitri, N.; Barone, V. Analytical First and Second Derivatives for a Fully Polarizable QM/Classical Hamiltonian. *J. Chem. Theory Comput.* **2012**, *8*, 4270–4278.
- (17) Lipparini, F.; Cappelli, C.; Barone, V. A Gauge Invariant Multiscale Approach to Magnetic Spectroscopies in Condensed Phase: General Three-Layer Model, Computational Implementation and Pilot Applications. *J. Chem. Phys.* **2013**, *138*, 234108.
- (18) Mennucci, B. Polarizable Continuum Model. *WIREs Comput. Mol. Sci.* **2012**, *2*, 386–404.
- (19) Klamt, A. The COSMO and COSMO–RS Solvation Models. *WIREs Comput. Mol. Sci.* **2011**, *1*, 699–709.
- (20) Greengard, L.; Rokhlin, V. A Fast Algorithm for Particle Simulations. *J. Comput. Phys.* **1987**, *73*, 325–348.
- (21) Scalmani, G.; Barone, V.; Kudin, K.; Pomelli, C.; Scuseria, G.; Frisch, M. Achieving Linear-Scaling Computational Cost for the Polarizable Continuum Model of Solvation. *Theor. Chem. Acc.* **2004**, *111*, 90–100.
- (22) Lipparini, F.; Scalmani, G.; Mennucci, B.; Cancès, E.; Caricato, M.; Frisch, M. J. A Variational Formulation of the Polarizable Continuum Model. *J. Chem. Phys.* **2010**, *133*, 014106.
- (23) Lipparini, F.; Scalmani, G.; Mennucci, B.; Frisch, M. J. Self-Consistent Field and Polarizable Continuum Model: A New Strategy of Solution for the Coupled Equations. *J. Chem. Theory Comput.* **2011**, *7*, 610–617.
- (24) York, D.; Karplus, M. A Smooth Solvation Potential Based on the Conductor-Like Screening Model. *J. Phys. Chem. A* **1999**, *103*, 11060–11079.
- (25) Scalmani, G.; Frisch, M. J. Continuous Surface Charge Polarizable Continuum Models of Solvation. I. General Formalism. *J. Chem. Phys.* **2010**, *132*, 114110.
- (26) Lange, A. W.; Herbert, J. M. Polarizable Continuum Reaction-Field Solvation Models Affording Smooth Potential Energy Surfaces. *J. Phys. Chem. Lett.* **2010**, *1*, 556–561.
- (27) Lange, A. W.; Herbert, J. M. A Smooth, Nonsingular, and Faithful Discretization Scheme for Polarizable Continuum Models: The Switching/Gaussian Approach. *J. Chem. Phys.* **2010**, *133*, 244111.
- (28) Strain, M. C.; Scuseria, G. E.; Frisch, M. J. Achieving Linear Scaling for the Electronic Quantum Coulomb Problem. *Science* **1996**, *271*, 51–53.
- (29) Cancès, E.; Maday, Y.; Stamm, B. Domain Decomposition for Implicit Solvation Models. *J. Chem. Phys.* **2013**, *139*, 054111.
- (30) Lipparini, F.; Stamm, B.; Cancès, E.; Maday, Y.; Mennucci, B. Fast Domain Decomposition Algorithm for Continuum Solvation Models: Energy and First Derivatives. *J. Chem. Theory Comput.* **2013**, *9*, 3637–3648.
- (31) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. Development and Use of Quantum Mechanical Molecular Models. 76. AM1: A New General Purpose Quantum Mechanical Molecular Model. *J. Am. Chem. Soc.* **1985**, *107*, 3902–3909.
- (32) Stewart, J. J. P. Optimization of Parameters for Semiempirical Methods. 1. Method. *J. Comput. Chem.* **1989**, *10*, 209–220.
- (33) Stewart, J. J. P. Optimization of Parameters for Semiempirical Methods V: Modification of NDDO Approximations and Application to 70 Elements. *J. Mol. Model.* **2007**, *13*, 1173–1213.
- (34) Ridley, J.; Zerner, M. An Intermediate Neglect of Differential Overlap Technique for Spectroscopy: Pyrrole and the Azines. *Theor. Chem. Acc.* **1973**, *32*, 111–134.
- (35) Zerner, M. C. *Reviews in Computational Chemistry*; John Wiley & Sons, Inc.: New York, 2007; pp 313–365.
- (36) Khandogin, J.; Gregersen, B. A.; Thiel, W.; York, D. M. Smooth Solvation Method for d-Orbital Semiempirical Calculations of Biological Reactions. 1. Implementation. *J. Phys. Chem. B* **2005**, *109*, 9799–9809.
- (37) 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.; et al. *Gaussian Development Version*, revision H.33+; Gaussian Inc.: Wallingford, CT, 2010.
- (38) Hofmann, E.; Wrench, P. M.; Sharples, F. P.; Hiller, R. G.; Welte, W.; Diederichs, K. Structural Basis of Light Harvesting by Carotenoids: Peridinin-Chlorophyll-Protein from *Amphidinium carterae*. *Science* **1996**, *272*, 1788–1791.
- (39) Jordanides, X. J.; Lang, M. J.; Song, X.; Fleming, G. R. Solvation Dynamics in Protein Environments Studied by Photon Echo Spectroscopy. *J. Phys. Chem. B* **1999**, *103*, 7995–8005.
- (40) Kowalska, D.; Krajnik, B.; Olejnik, M.; Twardowska, M.; Czechowski, N.; Hofmann, E.; Mackowski, S. Carotenoid to Chlorophyll Energy Transfer in the Peridinin-Chlorophyll-A-Protein Complex Involves an Intramolecular Charge Transfer State. *Proc. Nat. Acad. Sci. U.S.A.* **2002**, *2013*, 16760–16765.