A polarizable mixed Hamiltonian model of electronic structure for micro-solvated excited states. I. Energy and gradients formulation and application to formaldehyde ($^1A_2$)

M. Dupuis[a)]
Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, Washington 99352 and Department of Applied Chemistry, Graduate School of Engineering, The University of Tokyo, Tokyo, Japan 113

M. Aida
Department of Chemistry, Hiroshima University, 1-3-1 Kagamiyama, Higashi-Hiroshima 739-8526, Japan

Y. Kawashima and K. Hirao
Department of Applied Chemistry, Graduate School of Engineering, The University of Tokyo, Tokyo, Japan 113

We describe an efficient implementation of a polarizable mixed Hamiltonian model of electronic structure that combines Hartree–Fock, Kohn–Sham, or multiconfiguration quantum-chemical wave functions with a polarizable and flexible molecular mechanics potential of water, and that is applicable to micro-solvated electronic excited states. We adopt a direct algorithm for the calculation of the polarization response of the solvent subsystem. The strategy facilitates the calculation of the energy of the system and of the forces with respect to the solute coordinates and the solvent coordinates, including for excited states. This capability opens the way to the determination of optimized, transition structures, force constants, and intrinsic reaction pathways for the solute–solvent system, and to molecular dynamics calculations to account for finite temperature effects. As an illustration we characterize the structure and energy of micro-solvated formaldehyde H_2CO in its ground state and in its $^1(\pi^*\leftarrow n)$ excited state. A novel perpendicular structure is found to be the lowest energy conformation of the $H_2CO\,^1(\pi^*\leftarrow n):H_2O$ complex. The all-quantum-chemical results and the mixed Hamiltonian results, with or without solvent polarizability, are in semiquantitative agreement. We comment on the choice of Lennard-Jones parameters associated with a solute excited state. Lennard-Jones parameters that yield good ground state structures and energies with the mixed Hamiltonian model, are found to be too soft for the micro-solvated excited state H_2CO in the adiabatic (equilibrium micro-solvation) regime. © *2002 American Institute of Physics.* [DOI: 10.1063/1.1483858]

I. INTRODUCTION

A. Continuum and molecular solvation models

Much effort is directed towards extending the array of electronic structure methods from the gas phase to the condensed phase.[1–4] Modeling of solvation can be broadly categorized along two types of models, the continuum models[1,2] and the discrete models,[3,4] with both types continuing to present challenges.

Early computational implementations of solvation effects used the dielectric continuum representation of the solvent, starting with Onsager's spherical cavity model in the context of semiempirical and *ab initio* wave functions.[5,6] Most recent continuum models involve a description of the solute residing inside a molecule-shaped cavity embedded in a dielectric continuum.[1,7–9] The electron density and nuclei of the solute polarize the continuum and induce partial charges on the cavity boundary. See Refs. 1 and 2 and references therein as excellent reviews of various modern con-

tinuum models and as a comprehensive source of information. A large body of calculated chemical data is already available from the use of these approaches. The continuum models have been found successful and practical for many types of molecular systems. Generally the cavity is constructed from interlocked spheres of given radii[1,9] or as a solute iso-density surface.[10–12] It is fair to say that the results are sensitive to the choice of the atomic radii or of the iso-density contour values, in particular for ionic solutes. This issue remains an area of active research[13,14] in particular to try to account for the detailed interaction and structure of the important (first) hydration shells around the solute in cases of strong solute–solvent interaction. To date, most continuum models have been used to describe equilibrium solvation, although methods to deal with nonequilibrium solvation and the fast electronic response of the solvent have progressively been developed[1,2,15–18] and practical formulations are available. Although continuum models have also been applied with satisfactory results to describe the solvation effects in transition state structures involving bond breaking and bond formation in several cases,[2,19–21] the selection of atomic radii

---

[a)]Electronic mail: michel.dupuis@pnl.gov

1242

for transition states is likely to remain an even more acute issue than for equilibrium structures.[21]

Discrete solvation models continue to be developed and are used increasingly.[3,4,22–40] Among these models the most natural model is the all-quantum-mechanical (QM) model of micro-solvation where water molecules are explicitly included in the quantum chemical description of the system. Such a model is well adapted to the description of the detailed structure of the solute–solvent interface near equilibrium and along reaction pathways. However the approach is strongly limited by the computational expense of the quantum-chemical level of theory and is restricted to small numbers of water molecules.[41] The most powerful incarnation of this approach is actually found in the application to solvation problems[42–44] of the Car–Parrinello *ab initio* molecular dynamics approach based on plane-wave density-functional theory[45] (DFT). Alternative approaches, denoted QM/MM, involve the representation of the solvent molecules by means of classical molecular mechanics (MM) models for solvent–solvent interactions and of mixed quantum-classical models for solute–solvent interactions. See Refs. 3, 4, and 22–39 for early and recent accounts of mixed Hamiltonian methodologies and applications. The solvent–solvent MM/MM and solute–solvent QM/MM interactions are traditionally broken down into electrostatic and van der Waals contributions. The description of the electrostatic interactions varies in sophistication between different models, but in general there are semiquantitatively accurate parameterizations of these intereactions in the many published MM/MM potentials. In QM/MM models the electrostatic representation of the electron–nuclei interaction with the MM solvent molecules is either done exactly, or accurately through distributed multipolar expansions. In some cases the solute–solvent interaction is accounted for by means of a mean potential from the solvent.[30–32] The van der Waals interactions describe the repulsion and the attractive dispersion interactions of the solute–solvent fragments as they come in proximity. The quality of the total QM/MM interactions requires a balanced description between the electrostatic and van der Waals interactions. Molecular mechanics MM potentials have been developed and used with success for many years to predict free energies of solvation and structures of macromolecules.

In most QM/MM models the interaction between the solute and the solvent molecules uses a dual description of the solute. The electrostatic interaction is described as the interaction between the electron density and the nuclei of the QM solute with the effective charges (and possibly the induced dipoles when using polarizable MM potentials) of the classical MM solvent model. The electron density is determined by means of a quantum chemical method, most often the Hartree–Fock or the Kohn–Sham wave function. The van der Waals interactions between the solute and the solvent molecules are calculated through a classical representation of the solute with a Lennard-Jones (LJ) potential function similar to the one used for solvent–solvent interaction. This dual representation of the solute, quantum for electrostatics and classical for the van der Waals interactions, is a possible source of inconsistencies of course. To make up for these inconsistencies the LJ parameters of the "pseudo" solute are sometimes adjusted.[33] To the best of our knowledge no LJ parameters have been proposed for excited-state solutes, although it is anticipated that the LJ diameter for excited states should be larger than for ground states, owing to the more diffuse nature of excited states.

In recent years classical force fields have been enhanced with the capability of describing the electronic polarization of the solvent by means of polarizable intermolecular potentials. Models with polarizable potentials[46–47] have been shown to give an improved description of liquids and of solute–solvent interactions. The ability to treat solvent polarization is also critical to describe non-adiabatic processes in solution, such as electronic excitation, electron transfer, and charge transfer processes. For these processes the electronic response of the solvent is much faster than its orientational (inertial) response, so that it is important that electronic solvent polarization be accounted for. With this in mind a few hybrid models have been developed that include solvent polarization.[36–40] Note that the effective fragment potential (EFP) model[40] includes an electronic polarization of the solvent molecules as a result of the form of the molecular EFP of the solvent. Relatively few molecular EFP's exist due to a certain complexity of their development. Among the QM/MM implementations the polarizable models are much less widely used, presumably owing to the current implementations (cumbersome integral expansions in the direct reaction field implementation,[27] and dual MM and QM self-consistency algorithm[36–38]).

B. An enhanced molecular model for micro-solvated excited states

We report here on a new, efficient implementation of a QM/MM-pol model in which the MM potential may include a polarization term. The formalism for the energy and gradient calculations of this hybrid QM/MM-pol model has been presented before by others.[27,29,36–38] Our contribution is a much simpler, more efficient, and more practical implementation of the QM/MM-pol model than reported heretofore. Additionally we extend the formalism and implementation to the calculation of excited states using the multiconfigurational Hartree–Fock (MCSCF) level of theory for the solute and to the calculation of the energy gradients, with respect to both the solute and the solvent. Note that gradient expressions have been reported for QM/MM and QM/MM-pol models[29,36,37] for ground-state solutes. Here we extend the formalism to the QM/MM-pol model for excited states and the forces acting on either the QM or the MM subsystem. The ability to calculate the QM and MM forces opens up the whole array of automatic tasks that make use of analytic gradients, such as determination of minimum energy and transition states structures, determination of intrinsic reaction coordinate pathways, as well as finite temperature molecular dynamics approaches for solvated systems.

One additional characteristic feature of the model used in the present work lies in the treatment of the vibrational motion of the solvent MM molecules. In most cases reported to date, QM/MM calculations assume rigid solvent molecules: The rigidity constraint is not an issue in Monte Carlo

treatments of the solvent. In molecular-dynamics simulations, specifically designed algorithms must be used to maintain the rigid solvent structure. In contrast, in the present work we follow the approach of Lie and Clementi[48] and superimpose an intramolecular vibrational potential to the MM intermolecular potential. Thus the solvent molecules have the ability to distort. We denote this model QM/MM-pol-vib to reflect that the MM subsystem is represented by a polarizable potential and the solvent water molecules are not constrained to be rigid. From a model point of view this approach permits one to include the effects of solvent vibrations. From an algorithmic point of view the approach has the advantage that the molecular-dynamics equations are simpler and it is easier to deal with the forces acting on the solvent molecules. In a brief preliminary account[49] we reported the application of this model to the characterization of the structure of cyclic water clusters up to pentamers using the Hartree–Fock level of theory for the QM subsystem. In that work we systematically replaced one QM water molecule at a time in a cluster with one MM molecule, and the mixed cluster was re-optimized. We found that the mixed cluster structures and energies reproduced well some of the finer details of the structures and energies of the all-QM clusters. One of the elements that we carefully watched was whether the distortions observed in the flexible MM molecules in the mixed clusters when using the QM/MM-pol-vib model reflected the structural features of the all-QM clusters. This was indeed the case. In the present work we essentially extend this earlier analysis to the structure of a solvated electronically excited solute.

We recently reported[50,51] a study of an S_N2 type II reaction, the hydrolysis of chloromethane, carried out by means of *ab initio* molecular-dynamics using HF/6-31G* energies and forces. The model system included several "solvent" water molecules treated at the QM level of theory in addition to the "active" water molecule that participated in the hydrolysis. It was found that in the transition state region of the reaction the solvent molecules adopt a networklike arrangement in solvating the active system. Furthermore it was observed that these solvent molecules are involved in processes of proton relays with nearly concerted proton transfers among the solvent molecules. These solvent molecules play an active role in the reaction and need to be treated at the quantum level of theory. Numerous other examples of the reactive role of solvent molecules in micro-solvated systems are being reported.[41–44] However, beyond the quantum treatment needed for these few solvent molecules, the boundary between the QM subsystem and the MM subsystem need to be accurately represented. The present paper addresses again this very issue, but specifically with regards to excited states.

Our early experience with optimizing QM/MM clusters and carrying out molecular-dynamics simulations of these clusters indicated that the classical MM water molecules not belonging to the inner solvation shells show a tendency to drift away. Thompson and Schenter[36] reported a similar experience and suggested superimposing a soft quadratic potential to prevent drifting. In the present work we describe a scheme whereby the solute–solvent cluster is enclosed in a spherical cavity whose wall interact with either the solute or the solvent molecules by means of a repulsive LJ potential. The global model is, therefore, labeled QM/MM-pol-vib/CAV. In the present implementation the outside of the spherical cavity is vacuum. In future reports we will describe a simple extension of the present model in which a dielectric continuum lies outside the cavity.

C. A case study: Solvation of excited formaldehyde H₂CO [$^1(\pi^*\leftarrow n)$]

As an illustration of the present development, we initiated a QM/MM-pol-vib/CAV study of the relaxation dynamics of "aqueous" formaldehyde following a $^1(\pi^*\leftarrow n)$ excitation as a prototype of carbonyl excitation dynamics.[52] This project brings up three stringent aspects of the solute–solvent description: (1) An accurate representation of the solute–water structures and energetics for the $H_2CO(^1A_2):H_2O$ system; (2) a satisfactory representation of the solute in an "aqueous" environment (i.e., a larger set of H_2O molecules); and (3) a statistical treatment of the solvent molecular environment (i.e., configuration sampling via a Monte Carlo or molecular-dynamics treatment, still in the context of high-level QM mixed with MM). In this and subsequent papers we systematically investigate these issues. In the present paper we investigate the ability of the new mixed Hamiltonian model to describe accurately the interaction of an electronically excited solute with a water molecule. In a companion paper[53] we report an investigation of the excitation blue shift in micro-solvated and aqueous formaldehyde. A computationally efficient statistical treatment of the solvation effects will be presented in a forthcoming publication.[54]

The present paper focuses solely on the structure and energetics of the complex $H_2CO(^1A_2):H_2O$ between electronically excited formaldehyde and one water molecule. The QM level of theory is of a multiconfiguration type used in conjunction with a medium-size and an extended basis set. We find that the mixed models QM/MM with a nonpolarizable potential and a polarizable potential of water are both in semiquantitative agreement with the results of all quantum-mechanical calculations. No significant differences are observed when using a polarizable MM potential or a nonpolarizable MM potential for the solvent water. This finding is in accord with the results from Bader *et al.*[55] obtained with all-MM models. We include also in this paper a preliminary quantitative assessment of the generally anticipated inaccuracies associated with using, for excited-state calculations, LJ parameters defined for ground-state systems.[55] We find that "ground state" LJ parameters lead to an underestimation of the van der Waals repulsion in the excited state in the adiabatic equilibrium regime. Preliminary results indicate that slightly larger values (by $\sim 5\%$) of the LJ diameter σ would yield improved agreement with all-QM interaction potentials.

The paper is organized as follows: In Sec. II we give on overview of the QM/MM-pol-vib/CAV implementation with the detailed equations given in Appendices A and B. In Sec. III we present and discuss the results concerning the prototypical system, electronically excited formaldehyde $H_2CO(^1A_2):H_2O$ micro-solvated by one water molecule. We summarize the present research in Sec. IV.
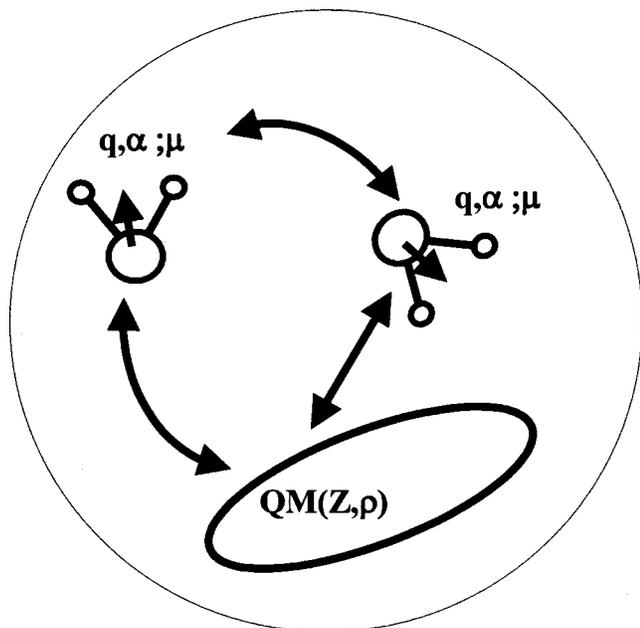
FIG. 1. Schematic representation of the QM and MM interactions included in the QM/MM-pol-vib formalism.

II. ENERGY AND GRADIENTS IN THE QM/MM-pol-vib/CAV MODEL OF ELECTRONIC STRUCTURE: OVERVIEW

The molecular interactions that are included in the QM/MM-pol-vib/CAV model are depicted in Fig. 1. The many-electron state of the quantum subsystem (QM) is characterized by a Hamiltonian operator which includes the electrostatic potential due to the effective point charges assigned to the atomic centers of the MM subsystem, and to the effective induces dipoles also assigned to the MM atomic centers when atomic (isotropic) polarizability tensors are a part of the MM interaction potentials. In turn the induced point dipoles are determined from the electric field at the position of the polarizability tensors. There are several contributions to the electric field at any MM atomic position, one from the QM nuclei and the QM electron density, one from the effective charges from the other MM fragments, and one from the induced point dipoles at the other MM fragment positions. There is also a contribution arising from the self-interaction energy term (see below). A breakdown of the energy contributions is shown in Fig. 2 for various levels of approximation of the QM/MM solvation models. Detailed formulas and equations for the various operators and their integrals over Gaussian basis functions are given in Appendices A and B.

In keeping closely with the above description, Thompson and Schenter[36,37] and Gao[38] adopted a strategy based on a dual self-consistency of the QM wave function and the MM polarizable response. Their strategy is depicted in Fig. 3. An initial guess to the set of MM induced dipoles is derived and the resulting electrostatic potential is included into the one-electron Hamiltonian of the QM subsystem. The QM wave function is then solved to self-consistency, at which point new MM induced dipoles are generated from the converged QM wave function, and the QM wave function is

The QM/MM-pol-vib/CAV Model:

QM/MM model:
- QM electron density / MM point charges
- QM 'ad hoc' Lennard-Jones / MM Lennard-Jones
- MM / MM point charges and Lennard-Jones

QM/MM-pol model = QM/MM +
- QM electron density / MM induced dipoles
  (induced dipoles =   QM charge field / MM polarizabilities +
  MM charge field / MM polarizabilities +
  MM ind. dip. field / MM polarizabilities)

QM/MM-pol-vib model = QM/MM-pol +
- MM intra-molecular vibrational potential

QM/MM-pol-vib/CAV model = QM/MM-pol-vib +
- Lennard-Jones wall

$$E = E_{qm} + E_{qm/mm} + E_{mm/mm} + E_{pol} + E_{vib} + E_{(qm+mm)/cav}$$

FIG. 2. Breakdown of the energy contributions in the QM/MM-pol-vib/CAV formalism.

calculated to self-consistency anew. The process is repeated till both the QM electron distribution and the MM response are together self-consistent. Both of these research groups report that convergence is attained within three or four iterations of the (QM+MM) dual self-consistency.

Our improved implementation is depicted in Fig. 4. We calculate the MM induced dipoles every few iterations (one or two or three) of the QM self-consistency solution using the current QM electronic wave function. We include the electrostatic potential from the effective charges and from the updated dipoles into the one-electron QM Hamiltonian. We use the modified Hamiltonian in the determination of the improved electron density that yields a new set of MM induced dipoles. In effect we remove the explicit MM self-consistency and converge the QM and the MM "wave functions" simultaneously. In practice we find that updating the induced dipoles during the QM self-consistent field (SCF) process does not slow down the QM convergence, compared to what the convergence would be with fixed induced dipoles or in the gas phase for that matter.

An extension of the QM/MM-pol model, of which we have given a preliminary account,[49] is denoted QM/MM-pol-vib in which an intramolecular vibrational potential[48] is added to the MM/MM and QM/MM interaction potential. In addition to providing a more realistic model of the solvent water molecules, for example, the nonrigid treatment of the solvent molecules leads to computational simplifications in

Thompson-Schenter's QM/MM-pol Dual Self-Consistency Algorithm:

1.) calculate induced dipoles from
   - QM electron density,
   - MM polarizabilities,
   - MM point charges,
   - MM induced dipoles
2.) add (MM point charges + induced dipoles) potential to 1E-Ham
3.) new QM electron density (HF, MCSCF, DFT)

4.) go to 3.)   (QM SCF)

5.) go to 1.)   (MM SCF)

⇒ System SCF = QM SCF + MM SCF

FIG. 3. Schematic representation of the dual QM and MM self-consistency algorithm of Thompson and Schenter (Ref. 36).
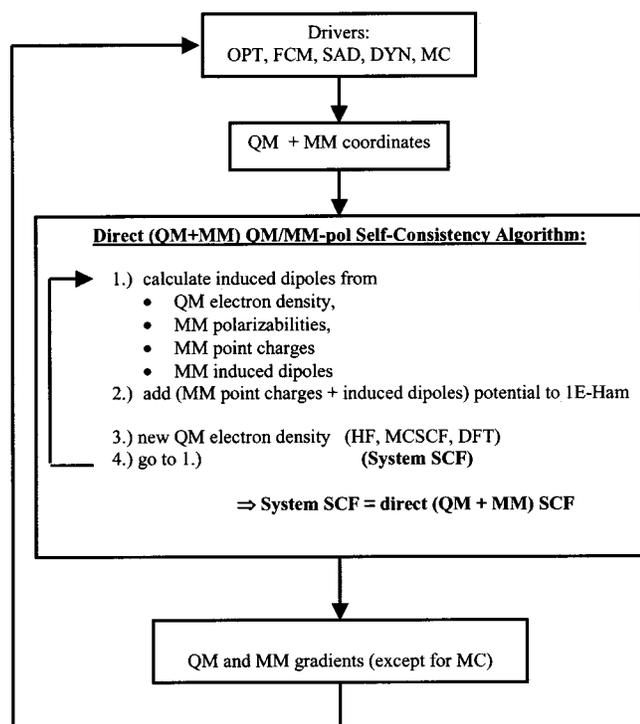
FIG. 4. Schematic representation of the direct (QM+MM) self-consistency algorithm for HF, KS, and MCSCF wave functions for ground and excited states. The configuration drivers are depicted: OPT=geometry optimization; FCM=force constant matrix calculation; SAD=transition state determination; DYN=molecular dynamics; MC=Monte Carlo sampling.

molecular dynamics calculations for example. Aida *et al.*[49] showed that such a treatment gives an accurate account of the solvation structure, as exemplified by the red shift in the frequency of the OH bond involved in hydrogen bonding.

An extension of the QM/MM-pol-vib model is a new model denoted QM/MM-pol-vib/CAV in which the total system is surrounded by a sphere which interacts with the total QM+MM system through a repulsive potential. As reported by Thompson[37] and observed by us as well, the MM water molecules have the tendency to drift away from the QM/MM system unless confined through a repulsive potential. Note that in the present model the sphere is not embedded in a

TABLE I. Lennard-Jones (LJ) Parameters for H_2CO And *tip3p* and *pol1* Parameters for H_2O.[a]

Molecule	H_2CO			H_2O	
Atom	O	C	H	O	H
σ[b]	3.600	3.800	2.600		
ε[b]	0.150	0.080	0.008		
q(*tip3p*)[c]				−0.834	0.417
σ(*tip3p*)				3.151	0.000
ε(*tip3p*)				0.152	0.000
q(*pol1*)[d]				−0.730	0.365
σ(*pol1*)				3.169	0.000
ε(*pol1*)				0.155	0.000
α(*pol1*)				0.465	0.135

[a](σ's are in Å, ε's are in kcal/mol, q's are in a.u., and α's are in Å^3).
[b]Parameters are taken from Freindorf and Gao, Ref. 33.
[c]*tip3p* parameters from Jorgensen, Ref. 61.
[d]*pol1* parameters from Dang, Ref. 46.

dielectric continuum as could be done to mimic long-range solvent effects. We would label such a model QM/MM-pol-vib/CONT. Development of such a model is in progress and will be reported in future accounts.

Another feature of our implementation lies in that we take advantage of a property of the MM polarization response. The MM induced dipoles minimize the energy of the total system with respect to changes in the induced dipoles (see Appendix A). This property leads to a significant simplication in the calculation of the energy gradients with respects to the Cartesian coordinates of both the QM and MM subsystems.[56] We carried out the above implementation for closed-shell and open-shell HF and KS wave functions as well as MCSCF wave functions. In the latter case the formalism applies to ground or excited states MC wave functions, as derived and illustrated below, with the sole difference between them coming from the calculation of the one-particle and two-particle density matrices of the QM wave function corresponding to either the ground state or to the excited state. The availability of analytic gradients permits the efficient determination of equilibrium and transition structures in equilibrium solvation as well as the ability to carry out classical molecular-dynamics calculations. The method has been implemented in the HONDO-2001 code.[57] Figure 4 depicts the structure of the program, how it is linked to the traditional drivers for geometry optimization, force constant calculation, saddle point determination, and also an *ab initio* molecular-dynamics driver[58] and a Monte Carlo solvent configuration generator.[54]

III. APPLICATION TO THE $H_2CO(^1A_2)$:H_2O COMPLEX

To illustrate the method and its implementation we report structures, energies, and vibrational frequencies of a hydrogen-bonded complex between the formaldehyde molecule in its singlet $^1(\pi^* \leftarrow n)$ excited state and a water molecule. These calculations are a part of a larger project aiming at characterizing the relaxation dynamics of formaldehyde in aqueous phase after vertical absorption, using an accurate QM level of theory mixed with a MM representation of the aqueous medium. The complex is denoted $H_2CO(^1A_2)$:*1*w to indicate that the formaldehyde moiety is in an excited state that has parentage in the $H_2CO(^1A_2)$ state of formaldehyde in the gas phase. The short-hand notation H_2CO:*1*w is applied at times. We used two basis sets, the 6-31G* basis set[59] and the cc-pVTZ (denoted *vtz*) basis set of Dunning.[60] For MM water potentials we used the nonpolarizable *tip3p* potential of Jorgensen[61] and the polarizable *pol1* potential of Dang.[46] The water vibrational potential was the one from Bartlett *et al.*[62] The QM wave function was a multiconfiguration wave function of the complete active space (CAS) type[63] that includes all the electronic configurations obtained by distributing all 12 valence electrons of H_2CO among 10 valence orbitals of H_2CO consistent with spin and space symmetry. The excited-state wave functions was obtained by optimizing the orbitals for the second root of the CAS configuration interaction expansion in C_1 symmetry or the first root in C_s symmetry. The QM/MM wave functions are labeled $CAS(^1A_2)$/6-31G*:*tip3p*, $CAS(^1A_2)$/6-31G*:*pol1*, $CAS(^1A_2)$/vtz:*tip3p*, and $CAS(^1A_2)$/vtz:*pol1*. In the
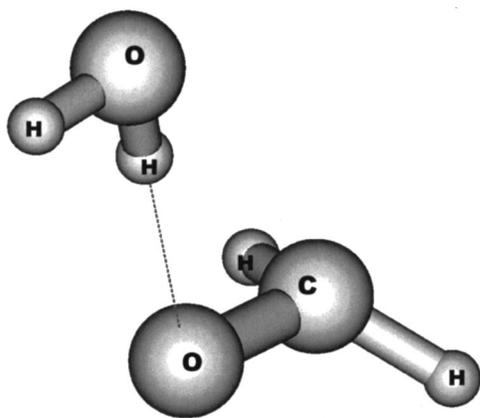
FIG. 5. The lowest energy structure of the $H_2CO(^1A_2){:}H_2O$ molecular system.

QM/MM calculations we used the LJ parameters suggested by Freindorf and Gao[33] for formaldehyde, even though the suggested parameters were determined on the basis of ground state calculations for a variety of "solute:1w" systems. We comment below on the validity of these "ground-state" LJ parameters for excited-state calculations. For completeness the LJ parameters are gathered in Table I. For comparison purposes we calculated also all-QM wave functions denoted $CAS(^1A_2)/6\text{-}31G^*{:}6\text{-}31G^*$ and $CAS(^1A_2)/vtz{:}vtz$ with the water molecule treated quantum mechanically and its 10 electrons occupying five molecular orbitals that are essentially localized on water. A picture of the complex is shown in Fig. 5 and the calculated structural, energetic, and vibrational data are given in Tables II and III.

The most notable feature of the $H_2CO(^1A_2){:}H_2O$ complex is that it has a perpendicular conformation, with the water molecule residing in the plane bisecting the formaldehyde molecule and containing the C–O bond. The water molecule acts as a hydrogen-bond donor. The three levels of theory yield structures of the $H_2CO(^1A_2)$ moiety that are very similar, a C–O bond length of ~ 1.361 Å and an out-of-plane angle $\delta(OC\text{–}H_2) \sim 38.4$ degrees, which are very close to the isolated molecule values (1.359 Å and 38.7 deg.). The OH bond of water that is involved in the hydrogen-bond is slightly elongated in all three models, the elongations in the QM/MM(-pol) calculations being very similar to the elongation in the all-QM calculation. This finding is consistent with our earlier report on mixed QM/MM water clusters.

The harmonic frequencies of the complexes are given in Table III. The modes can be assigned a type depending on whether they involve mostly the water moiety, the formaldehyde moiety, or whether they are intermolecular modes. The modes assigned to the H_2CO moiety have frequencies close to those of an isolated $H_2CO(^1A_2)$. The modes assigned to the water moiety are different for the CAS:tip and CAS:pol models compared to the all-QM model. This is a reflection of the different vibrational force field of the water molecules in the complex, the Hartree–Fock force field in the all-QM calculation vs the highly accurate configuration interaction (CI)-derived force field in the QM/MM(-pol) calculations. The

inter-molecular modes are those most relevant to the description of the complex. There is overall very good agreement between all the models. The largest frequency mode among those at ~ 390 cm^{-1} corresponds to the hydrogen bond stretching mode with the CAS/tip description being slightly stiffer than the CAS/pol and the all-QM descriptions, this in spite of the hydrogen bond lengths showing a large disparity.

The hydrogen bond energies (E_{HB}) obtained with the 3 levels of theory are in semiquantitative agreement, 2.1, 2.7, and 2.4 kcal/mol for the $CAS(^1A_2)/6\text{-}31G^*{:}6\text{-}31G^*$, $CAS(^1A_2)/6\text{-}31G^*{:}tip3p$, and $CAS(^1A_2)/6\text{-}31G^*/6\text{-}31G^*{:}pol1$ respectively, and nearly the same with the vtz basis set. The all-QM values of E_{HB} listed in the table were corrected for basis set superposition error.[64]

The hydrogen-bond length $R(H\cdots O)$, the hydrogen-bond angles $\alpha(H\cdots O\text{–}C)$ and $\alpha(O\text{-}H\cdots O)$, and the hydrogen-bond torsion angle $\tau(O\text{–}H\cdots O\text{–}C)$ are the most relevant geometrical parameters in assessing the accuracy of the models. The accord between the all-QM and mixed QM/MM models is semi-quantitative. The most interesting parameter is the hydrogen-bond length $R(H\cdots O)$ for which there are significant structural differences between all three models, albeit the binding energies are very similar. This distance is a reflection of the proper balance, or lack of such, between the electrostatic contribution to the bond energy and the long-range repulsion–attraction of the LJ term that enters the QM/MM-pol-vib model. The model uses an *ad hoc* LJ (6-12) potential between the QM subsystem and the MM subsystem. For the MM subsystem we used the $tip3p$ and $pol1$ LJ parameters while for the QM subsystem we use the modified LJ parameters suggested by Friendorf and Gao.[33] These later parameters were determined on the basis of a number of solute–water structures in their ground states. It is generally recognized that the LJ parameters for an excited state solute–water system should be different, albeit we are not aware of published values for excited states. Bader *et al.*[55] argued that larger values of the LJ parameters for the solute would be appropriate, considering the larger size of the orbitals and polarizability of the formaldehyde excited state compared to its ground state. (However, these authors followed Levy *et al.*[52] in using the same LJ parameters for the excited state as for the ground state.) In the present work we also used the same "ground-state" parameters. An all-QM calculation of the same type for ground-state formaldehyde $CAS(^1A_1)/6\text{-}31G^*{:}6\text{-}31G^*$ yielded a $R(H\cdots O) \sim 2.112$ Å, and $CAS(^1A_1)/6\text{-}31G^*{:}tip3p$ and $CAS(^1A_1)/6\text{-}31G^*{:}pol1$ values of ~ 2.097 and ~ 2.098 Å, respectively (see companion paper[53]). For the ground state the discrepancy between the all-QM and the QM/MM models amounts to ~ 0.02 Å. The discrepancy of 0.10 and 0.07 Å for the QM:$tip3p$ and QM:$pol1$ models with the $O\cdots H$ bond shorter than for the all-QM structure, is much larger for the excited state than for the ground state. The same conclusions hold for the calculations using the vtz basis set. Thus it appears indeed that the LJ parameters for electronically excited formaldehyde need to be larger than for ground-state formaldehyde, presumably to increase the repulsion interaction between solute and water. That this may be the case is further supported by the $pol1$ results. It can be seen from

TABLE II. Structure of the $H_2CO(^1A_2):H_2O$ complex for various all-QM and QM/MM(-pol) models.[a]

	CAS/6-31G*: 6-31G*	CAS/6-31G*: tip3p	CAS/6-31G*: pol1	CAS/vtz: vtz	CAS/vtz: tip3p	CAS/vtz: pol1
R(C–H)	1.106;1.106	1.106;1.106	1.106;1.106	1.104;1.104	1.104;1.104	1.104;1.104
α(HCO)	114.0;114.0	114.0;114.0	114.0;114.0	114.5;114.5	114.4;114.4	114.4;114.4
R(C–O)	1.363	1.362	1.362	1.357	1.358	1.357
R(H$\cdots$O)	2.246	2.150	2.178	2.293	2.175	2.197
α(H$\cdots$O–C)	112.8	113.7	113.3	113.9	113.7	113.4
α(O–H$\cdots$O)	179.4	173.9	173.2	179.0	171.9	171.7
τ(O–H$\cdots$O–C)	0.0	180.0	180.0	0.0	180.0	180.0
R(O–H)[b]	0.949;0.947	0.964;0.958	0.963;0.958	0.942;0.941	0.963;0.958	0.963;0.958
α(H–O–H)	105.4	103.6	103.7	105.8	103.6	103.7
τ(H–O–H$\cdots$O)	180.0	0.0	0.0	180.0	0.0	0.0
δ(OC–H$_2$)	38.6	38.6	38.6	37.1	37.2	37.2
Energy(a.u.)	−189.872 074	−113.861 314	−113.860 707	−189.957 493	−113.902 108	−113.901 520
E_{HB}(kcal/mol)[c]	2.1	2.7	2.4	1.8	2.6	2.3

[a]Bond lengths in Å, angles in degrees, energies in a.u., hydrogen-bond energies E_{HB} are in kcal/mol.
[b]The first value is for the H atom involved in the hydrogen bond.
[c]The binding energy E_{HB} was calculated using the energy of the complex at infinite distance treated as a supermolecule and is corrected for basis set superposition error (BSSE) for the all-QM calculations. The CAS/6-31G* energy of isolated $H_2CO(^1A_2)$ is −113.856 927 a.u. The CAS/vtz energy of isolated $H_2CO(^1A_2)$ is −113.897 894 a.u.

Table I that the LJ parameters are essentially the same for *pol1* as for *tip3p* while the partial charges are smaller in *pol1* than in *tip3p*. The fact that the H$\cdots$O distance in the CAS($^1A_1$)/6-31G*:*pol1* complex is longer than in the CAS($^1A_1$)/6-31G*:*tip3p* complex while the LJ energy contribution is nearly the same, suggests that the electrostatic contribution to the binding energy is less with *pol1* than with *tip3p*, in accord with the reduced partial charges, even though the atomic polarizabilities contribute also to the electrostatic interaction binding energy. The fact remains that the solute–water attraction is too strong and that a stronger repulsion is needed to obtain better agreement between the QM/MM models and the all-QM data.

In support for this finding we obtained the optimized structures for the $H_2CO(^1A_2):H_2O$ complex with modified LJ parameters for the oxygen atom of formaldehyde. In this preliminary investigation we selected to modify the LJ parameters for the carbonyl oxygen only because changes in the oxygen parameters are likely to have the largest effects on the QM/MM(-pol) structure of the complex, in particular on the hydrogen-bond length. There are differences in angles as well between the all-QM and the QM/MM(-pol) structures. Adjusting all the parameters would be needed to obtain close agreement, along the lines of the work of Freindorf and Gao.[33] A full investigation is beyond the scope of this work albeit it is in progress. More specifically we increased the value of $\sigma(O)$ by 5% and 10%, while holding $\varepsilon(O)$ constant as well as the LJ parameters for C. The parameters for H_2O were not changed. The results are displayed in Table IV. It can be seen that the 5% increase yields a O$\cdots$H bond length

TABLE III. Harmonic frequencies of the $H_2CO(^1A_2):H_2O$ complex for various all-QM and QM/MM(-pol) models.[a]

Freq. (cm^{-1})	Sym[a]	Type[b]	CAS/6-31G*: 6-31G*	CAS/6-31G*: tip3p	CAS/6-31G*: pol1	CAS/vtz: vtz	CAS/vtz: tip3p	CAS/vtz: pol1
ω_1	a′	w-f	48	51	47	47	51	47
ω_2	a″	w-f	73	61	63	55	66	55
ω_3	a″	w-f	104	115	108	97	113	106
ω_4	a′	w-f	116	136	132	105	132	129
ω_5	a′	w-f	223	233	216	197	223	205
ω_6	a″	w-f	389	411	381	366	401	372
ω_7	a′	f	805	806	807	768	768	770
ω_8	a″	f	1001	999	999	990	989	989
ω_9	a′	f	1141	1143	1144	1126	1125	1127
ω_{10}	a′	f	1386	1385	1385	1368	1367	1367
ω_{11}	a′	w	1842	1704	1701	1778	1702	1700
ω_{12}	a′	f	2964	2963	2962	2921	2919	2919
ω_{13}	a″	f	3060	3057	3056	3016	3014	3013
ω_{14}	a′	w	4069	3782	3793	4122	3788	3797
ω_{15}	a′	w	4182	3923	3925	4215	3924	3927

[a]Symmetry group is C_s, the plane of symmetry contains the water molecule and bisects the formaldehyde molecule. See Fig. 5.
[b]Notation is: w for modes involving mostly the water molecule; f for modes involving the formaldehyde molecule; w-f for intermolecular modes.

TABLE IV. Structure of the $H_2CO(^1A_2):H_2O$ complex for various values of the LJ diameter for the carbonyl oxygen.[a]

	CAS/6-31G*: 6-31G*	CAS/6-31G*: tip3p	CAS/6-31G*: tip3p	CAS/6-31G*: tip3p
LJ for O:	N/A	Gao[b]	Gao+5%	Gao+10%
$\sigma(O);\varepsilon(O)$		3.60; 0.15	3.78; 0.15	3.96; 0.15
R(C–H)	1.106;1.106	1.106;1.106	1.106;1.106	1.106;1.106
α(HCO)	114.0;114.0	114.0;114.0	114.0;114.0	114.0;114.0
R(C–O)	1.363	1.362	1.362	1.362
R(H$\cdots$O)	2.246	2.150	2.247	2.353
α(H$\cdots$O–C)	112.8	113.7	112.3	111.2
α(O–H$\cdots$O)	179.4	173.9	172.9	172.4
τ(O–H$\cdots$O–C)	0.0	180.0	180.0	180.0
R(O–H)[b]	0.949;0.947	0.964;0.958	0.963;0.958	0.962;0.958
α(H–O–H)	105.4	103.6	103.7	103.7
τ(H–O–H$\cdots$O)	180.0	0.0	0.0	0.0
δ(OC–H$_2$)	38.6	38.6	38.6	38.6
Energy(a.u.)	−189.872 074	−113.861 314	−113.860 831	−113.860 436
E_{HB}(kcal/mol)	2.14	2.75	2.45	2.20

[a]Bond lengths in Å, angles in degrees, energies in a.u., hydrogen-bond energies E_{HB} are in kcal/mol.
[b]Parameters σ and ε taken from Freindorf and Gao (Ref. 33).
[c]The first value is for the H atom involved in the hydrogen bond.

in very good accord with the all-QM structure.

A proper description of the low-energy structure of $H_2CO(^1A_2):H_2O$ is an encouraging reflection of the ability of the model to describe well equilibrium solvation effects, including fluorescence. A second critically important criterion for the choice of proper LJ parameters for excited states lies in the ability to quantitatively describe the slope of the excited potential energy surface in the Franck–Condon region from the ground state. This attribute corresponds to the nonequilibrium regime of solvation. The present results indicate that the polarizable character of the *pol1* potential makes only a small contribution to the hydrogen bond description in so far as equilibrium "solvation" is concerned. A more extended investigation of the determination of proper LJ parameters for nonequilibrium solvation, and of the effect of polarizability in nonequilibrium solvation will be presented in a future report.[65]

IV. SUMMARY

In this paper we have presented a detailed description of an efficient implementation of a QM/MM-pol-vib/CAV model of micro-solvation with which we can calculate the energy and the forces of a QM subsystem embedded in an MM subsystem. The model includes a polarizable representation of the MM subsystem, and is equally applicable to the ground state or excited state of the QM subsystem. We carry out the implementation for the HF, DFT, and MCSCF levels of QM theory. The availability of analytic gradients for both the QM susbsystem and the MM subsystem makes it possible to apply efficient methods for the determination of equilibrium and transition structures, force constants, reaction pathways, and finite temperature dynamics trajectories.

We illustrated the capabilities of the method with a characterization of the structure and energy for the electronically excited formaldehyde–water system. We found that the all-QM and QM/MM structures and energies are in semi-quantitative agreement. We found that the LJ parameters in the QM/MM models used for the ground state give rise to too short hydrogen-bond lengths for the excited state. A preliminary study of LJ parameters for the excited state of formaldehyde suggests that the LJ diameter ought to be increased by ~5% for improved agreement between all-QM and QM/MM structures. An initial investigation of the solvation effects on the vertical absorption energy of formaldehyde as a prototypical carbonyl system is presented in a companion paper.[53] A full QM/MM-pol characterization of the dynamics of relaxation of excited formaldehyde will be forth coming.

APPENDIX A: QM/MM-pol-vib/CAV ENERGY

The total energy of a system in the QM/MM-pol-vib/CAV model has the form

$$E^{QM/MM} = E_{qm} + E_{qm/mm} + E_{mm/mm} + E_{pol} + E_{vib} + E_{(qm+mm)/cav}. \quad (A1)$$

E_{qm} includes to the traditional QM energy of the QM subsystem including the electrostatic interaction energy of the QM subsystem with the MM effective charges and dipoles. E_{qm-mm} refers to the LJ repulsion–dispersion interaction between the QM subsystem and the MM subsystem. $E_{mm/mm}$ refers to the electrostatic and LJ repulsion–dispersion interaction amongst the MM fragments. E_{pol} is the polarization energy limited to the MM induced dipoles among them-

selves. E_{vib} is the vibrational energy of the water solvent molecules arising from the nonrigid representation of the MM molecules. Finally $E_{(qm+mm)/cav}$ includes the LJ repulsive interaction between the QM and MM subsystems and the cavity wall. The various terms are defined and written out below.

1. The E_{qm} term

E_{qm} is the electrostatic interaction energy of the QM subsystem in the presence of the effective charges and of the induced dipoles of the MM subsystem. As such E_{qm} includes the electron–electron and the electron–nuclei energy terms within the QM subsystem, along with the interaction between the nuclei and the electron density with the MM effective charges and the interaction between the nuclei and the electron density with the MM effective dipoles. The potential acting on the electron density due to both the MM effective charges and the MM induced dipoles is included in the one-electron Hamiltonian operator for the QM subsystem. In what follows we give the equations for the most general case of a multiconfiguration self-consistent-field (MCSCF) wave function. Simplifying expressions for Hartree–Fock (HF) and Kohn–Sham (KS) wave functions are provided at the end of this section.

The QM wave function Ψ^S has the form of a linear combination of configuration state functions or equivalently in terms of Slater determinants $\{\Phi_I\}$, with superscript "S" denoting the "state," i.e., the ground state or an excited state

$$\Psi^S = \sum_I C_I^S \Phi_I, \tag{A2}$$

with Φ_I constructed from molecular orbitals ϕ_i expanded in terms of basis functions χ_μ

$$\phi_i = \sum_\mu c_{\mu i} \chi_\mu. \tag{A3}$$

The $C's$ are a solution of the configuration interaction secular equation, either for the ground state or an excited state

$$HC^S = EC^S, \tag{A4}$$

with

$$H_{IJ} = \sum_{ij} \gamma_{IJ}^{ij}(i|h|j) + \frac{1}{2}\sum_{ijkl} \Gamma_{IJ}^{ijkl}(ij|kl), \tag{A5}$$

and

$$E = \sum_{IJ} C_I^S H_{IJ} C_J^S. \tag{A6}$$

E takes the form

$$E = \sum_{ij} \gamma^{ij}(i|h|j) + \frac{1}{2}\sum_{ijkl} \Gamma^{ijkl}(ij|kl). \tag{A7}$$

γ^{ij} and Γ^{ijkl} are the one- and two-particle density matrix elements for state S, with

$$\gamma^{ij} = \sum_{IJ} C_I^S \gamma_{IJ}^{ij} C_J^S, \tag{A8}$$

$$\Gamma^{ijkl} = \sum_{IJ} C_I^S \Gamma_{IJ}^{ijkl} C_J^S. \tag{A9}$$

The ground state corresponds to the lowest eigenstate of the CI matrix [Eq. (A4)]. Higher energy eigenstates are the excited states. If "a" and "b" denote any doubly occupied orbital, "i" any variably occupied orbital, "x" any unoccupied orbital, and "p," "q," "r," and "s" any type of orbital, we have

$$\begin{cases} \gamma^{ab} = 2\delta_{a,a} \\ \gamma^{ai} = \gamma^{ax} = \gamma^{ix} = 0 \\ \Gamma^{pqrs} = \gamma^{pq}\gamma^{rs} - \frac{1}{4}(\gamma^{pr}\gamma^{qs} + \gamma^{ps}\gamma^{qr}) + \Gamma'^{pqrs}. \end{cases} \tag{A10}$$

Γ'^{pqrs} is nonzero only when all four orbitals are of the variably occupied orbital type (also called valence orbitals). In deriving the self-consistency equations for the MCSCF energy, we define the Fock operators and Lagrangian operators as follows:

$$F^{ij} = h\gamma^{ij} + \sum_{kl} |kl)\Gamma^{ijkl}, \tag{A11}$$

$$\varepsilon^{ri} = \sum_j (r|F^{ij}|j). \tag{A12}$$

We have

$$\varepsilon^{ri} = 2(F_{ri}^{core} + F_{ri}^{val}) \qquad \text{(Lagrangian, } i \text{ core),} \tag{A13}$$

$$\varepsilon^{ri} = \sum_j^{val} \gamma^{ij}F_{rj}^{core} + \sum_{jkl}^{val} \Gamma^{ijkl}(rj|kl)$$

$$\text{(Lagrangian, } i \text{ valence),} \tag{A14}$$

with F^{core} and F^{val} given by

$$F_{\mu\nu}^{core} = (\mu|h|\nu) + \sum_{\lambda\sigma} P_{\lambda\sigma}^{core}[(\mu\nu|\lambda\sigma) - \frac{1}{2}(\mu\lambda|\nu\sigma)], \tag{A15}$$

$$F_{\mu\nu}^{val} = \sum_{\lambda\sigma} P_{\lambda\sigma}^{val}[(\mu\nu|\lambda\sigma) - \frac{1}{2}(\mu\lambda|\nu\sigma)], \tag{A16}$$

with

$$P_{\lambda\sigma}^{core} = \sum_k^{core} 2c_{\lambda k}c_{\sigma k}, \tag{A17}$$

$$P_{\lambda\sigma}^{val} = \sum_{kl}^{val} \gamma^{kl}c_{\lambda k}c_{\sigma l}. \tag{A18}$$

We transform the particle density matrix elements to the AO basis following:

$$\gamma_{\mu\nu} = \sum_{ij} c_{\mu i}\gamma^{ij}c_{\nu j}, \tag{A19}$$

$$\Gamma_{\mu\nu\lambda\sigma} = \sum_{ijkl} c_{\mu i}c_{\nu j}c_{\lambda k}c_{\sigma l}\Gamma_{IJ}^{ijkl}. \tag{A20}$$

This leads to the energy expression

$$E = \sum_{\mu\nu} \gamma_{\mu\nu}(\mu|h|\nu) + \tfrac{1}{2}\sum_{\mu\nu\lambda\sigma}\Gamma_{\mu\nu\lambda\sigma}(\mu\nu|\lambda\sigma). \qquad (A21)$$

The Fock operators of Eqs. (A11), (A15), and (A16) play an important role in the various MCSCF convergence algorithms.[63]

2. HF and KS cases

In the case of an HF wave function, Eq. (A15) takes the forms

$$F_{\mu\nu} = (\mu|h|\nu) + \sum_{\lambda\sigma} P_{\lambda\sigma}[(\mu\nu|\lambda\sigma) - \tfrac{1}{2}(\mu\lambda|\nu\sigma)].$$
$$(A22)$$

$P_{\mu\nu}$ is the total density matrix element. The HF SCF equations for the molecular orbital coefficients c have the form:

$$Fc = Sc\varepsilon, \qquad (A23)$$

where S is the overlap matrix, $S_{\mu\nu} = (\mu|\nu)$, and ε represents the orbital eigenvalues. It follows that

$$E = \tfrac{1}{2}\sum_{\mu\nu} P_{\mu\nu}[(\mu|h|\nu) + F_{\mu\nu}]. \qquad (A24)$$

In the case of a KS wave function, we have

$$F_{\mu\nu} \equiv KS_{\mu\nu} = (\mu|h|\nu) + \sum_{\lambda\sigma} P_{\mu\nu}[(\mu\nu|\lambda\sigma)] + F_{\mu\nu}^{XC},$$
$$(A25)$$

with

$$F_{\mu\nu}^{XC} = \int \left(\frac{\partial F_{XC}}{\partial\rho}\chi_\mu\chi_\nu + \frac{1}{|\nabla\rho|}\frac{\partial F_{XC}}{\partial|\nabla\rho|}\nabla\rho\cdot\nabla(\chi_\mu\chi_\nu) \right)d^3r.$$
$$(A26)$$

In Eq. (A26) F^{XC} is the matrix of the Exchange-Correlation functional.[66] It follows in the HF case that:

$$E = \tfrac{1}{2}\sum_{\mu\nu} P_{\mu\nu}[(\mu|h|\nu) + F_{\mu\nu}]. \qquad (A27)$$

Solvation of the QM subsystem by the MM subsystem is reflected in the definition of the one-electron Hamiltonian $h(1)$ that appears throughout Eqs. (A1)–(A27). In the simple form of the QM/MM model a one-electron potential operator due to the MM effective charges (efc), labeled Q_s where the subscript s represents the MM atoms, is added to the usual kinetic energy and nuclear attraction operators. In the QM/MM-pol model, (isotropic) polarizabilities $\tilde{\alpha}_s$ are assigned to selected MM centers (often the MM atoms), and the electric field due to the QM subsystem as well as the other MM fragments induces point dipoles denoted $\mathbf{D}_s$ (efd). The kinetic energy and potential operators have the form

$$T(1) = -\tfrac{1}{2}\nabla^2(1), \qquad (A28)$$

$$V^{\mathrm{atm}}(1) = -\sum_A^{\mathrm{Nuc}} \frac{1}{r_{1A}} Z_A, \qquad (A29)$$

$$V^{\mathrm{efc}}(1) = -\sum_S^{N\mathrm{sol}} \frac{1}{r_{1S}} Q_S, \qquad (A30)$$

$$V^{\mathrm{efd}}(1) = -\sum_S^{N\mathrm{sol}} \nabla_S\!\left(\frac{1}{r_{1S}}\right)\cdot\mathbf{D}_S = -\sum_S^{N\mathrm{sol}} \frac{(\mathbf{r}_1 - \mathbf{R}_S)}{r_{1S}^3}\cdot\mathbf{D}_S,$$
$$(A31)$$

"Nuc" and "Nsol" throughout represent the number of atoms in the QM subsystem and in the MM subsystem, respectively. The minus sign in Eqs. (A29)–(A31) arises from the charge of the electron. Thus we have

$$h^0(1) = T(1) + V^{\mathrm{atm}}(1), \qquad (A32)$$

$$h(1) = h^0(1) + V^{\mathrm{efc}}(1) + V^{\mathrm{efd}}(1), \qquad (A33)$$

and

$$h_{\mu\nu} = \int \chi_\mu(1)h(1)\chi_\nu(1)d\tau_1. \qquad (A34)$$

It is convenient to define several energy terms involving the interaction of the electron density

$$V_{\mu\nu}^{\mathrm{atm}} = \int \chi_\mu(1)V^{\mathrm{atm}}(1)\chi_\nu(1)d\tau_1, \qquad (A35)$$

$$V_{\mu\nu}^{\mathrm{efc}} = \int \chi_\mu(1)V^{\mathrm{efc}}(1)\chi_\nu(1)d\tau_1, \qquad (A36)$$

$$V_{\mu\nu}^{\mathrm{efd}} = \int \chi_\mu(1)V^{\mathrm{efd}}(1)\chi_\nu(1)d\tau_1, \qquad (A37)$$

$$E^{\mathrm{elec/atm}} = \sum_{\mu\nu} P_{\mu\nu}V_{\mu\nu}^{\mathrm{atm}}, \qquad (A38)$$

$$E^{\mathrm{elec/efc}} = \sum_{\mu\nu} P_{\mu\nu}V_{\mu\nu}^{\mathrm{efc}}, \qquad (A39)$$

$$E^{\mathrm{elec/efd}} = \sum_{\mu\nu} P_{\mu\nu}V_{\mu\nu}^{\mathrm{efd}}. \qquad (A40)$$

These energy terms are in fact included in the electronic energy expression of Eq. (A21). We use Gaussian basis functions as the $\chi's$ and evaluate the one-electron integrals in Eq. (A34) by means of the generalized Rys numerical quadrature method[67] that was devised for the calculation of electrostatic potential, field, and field gradients, as well as their derivatives with respect to the QM atomic centers and the MM atomic centers for the calculation of the forces acting on the QM and MM subsystems. The generalization of the quadrature relies on a general Gaussian transform for the $1/r$ operator and its derivatives, shown in Appendix C. The generalization is systematic and its implementation in a computer code follows very much the simpler subroutine for the calculation of the nuclear attraction integrals of Eq. (A35). Compared to the approach adopted by others[27,36,37] in their earlier implementation of the QM/MM-pol method, the generalized Rys quadrature avoids the numerical inaccuracies otherwise associated with integral expansions and representations of point dipoles by finite dipoles, and more importantly, it also provides accurate energy derivatives.

The contributions to the interaction energy of the QM nuclei among themselves and of the QM nuclei with the MM charges and dipoles are

1252   J. Chem. Phys., Vol. 117, No. 3, 15 July 2002

Dupuis *et al.*

$$E^{\text{atm-atm}} = \frac{1}{2} \sum_{A \neq B}^{\text{Nuc}} Z_A \frac{1}{R_{AB}} Z_B, \tag{A41}$$

$$E^{\text{atm-efc}} = \sum_A^{\text{Nuc}} \sum_S^{\text{Nsol}} Z_A \frac{1}{R_{AS}} Q_S, \tag{A42}$$

$$E^{\text{atm-efd}} = \sum_A^{\text{Nuc}} \sum_S^{\text{Nsol}} Z_A \nabla_S \left(\frac{1}{R_{AS}} \right) \cdot \mathbf{D}_S$$

$$= \sum_A^{\text{Nuc}} \sum_S^{\text{Nsol}} Z_A \frac{(\mathbf{R}_A - \mathbf{R}_S)}{R_{AS}^3} \cdot \mathbf{D}_S. \tag{A43}$$

Finally we have

$$E_{\text{qm}} = E + E^{\text{atm/atm}} + E^{\text{atm/efc}} + E^{\text{atm/efd}}. \tag{A44}$$

3. The $E_{\text{qm/mm}}$ term

$E_{\text{qm/mm}}$ refers only to the van der Waals repulsion term between the "pseudo" QM subsystem and the MM subsystem (the electrostatic interaction terms are included in E_{qm} above). In its present form this term is only an *ad hoc* expression of the exchange repulsion and dispersion between the QM and MM subsystems. It is given by

$$E_{\text{qm/mm}} \equiv E_{\text{qm/mm}}^{\text{vdW}} = \sum_{A=1}^{\text{Nuc}} \sum_{S=1}^{\text{Nsol}} 4\varepsilon_{AS} \left[\left(\frac{\sigma_{AS}}{R_{AS}} \right)^{12} - \left(\frac{\sigma_{AS}}{R_{AS}} \right)^6 \right], \tag{A45}$$

with

$$\varepsilon_{AS} = (\varepsilon_A \varepsilon_S)^{1/2}, \tag{A46}$$

$$\sigma_{AS} = (\sigma_A \sigma_S)^{1/2}. \tag{A47}$$

In Eqs. (A45)–(A47), the ε_s's [not to be confused with the orbital energies of Eq. (A22)] and σ_s's are the usual LJ parameters that comprise the classical potential for the MM subsystem, the ε_A's and σ_A's are the LJ parameters of the classical potential for the "pseudo" QM subsystem. The summations run over the quantum atoms A and all the solvent atoms S. The LJ parameters for the QM subsystem can be taken directly from tabulated classical potentials. Freinforf and Gao[33] suggested slight adjustments to such parameters to match better the structure and energy results for selected systems obtained at the *ab initio* HF/6-31G* level of theory.

4. The $E_{\text{mm/mm}}$ term

The energy of the MM subsystem $E_{\text{mm/mm}}$ contains the usual electrostatic interaction between the effective charges of the MM atoms and also a van der Waals repulsion term

$$E_{\text{mm/mm}} = E^{\text{efc/efc}} + E_{\text{mm/mm}}^{\text{vdW}}, \tag{A48}$$

with

$$E^{\text{efc/efc}} = \frac{1}{2} \sum_{S \neq T}^{\text{Nsol}} Q_S \frac{1}{R_{ST}} Q_T, \tag{A49}$$

$$E_{\text{mm/mm}}^{\text{vdW}} = \sum_{S \neq T}^{\text{Nsol}} 4\varepsilon_{ST} \left[\left(\frac{\sigma_{ST}}{R_{ST}} \right)^{12} - \left(\frac{\sigma_{ST}}{R_{ST}} \right)^6 \right], \tag{A50}$$

$$\varepsilon_{ST} = (\varepsilon_S \varepsilon_T)^{1/2}, \tag{A51}$$

$$\sigma_{ST} = (\sigma_S \sigma_T)^{1/2}. \tag{A52}$$

The summations in Eqs. (A49)–(A50) are indexed with the condition $S \neq T$ to indicate that they run over all the solvent atoms S and T with the restriction that S and T belong to different residues, i.e., water molecules.

5. The E_{pol} term

The polarization energy E_{pol} here refers to energy terms involving the MM subsystem only. Terms involving the MM polarization interactions with the QM subsystem are already included in E_{qm}. The induced dipoles are determined by the response equations of the system to minimize the total energy of the QM/MM system. The electric field at the position of the polarizability tensors is separated into a static field and a dynamic field. The static field is due to the electron density and nuclei of the QM subsystem and to the effective charges of the MM subsystem. The dynamic field acting on a given induced dipole is due to the other induced dipoles. Thus the induced dipoles depend on the induced dipoles themselves. The response equations are thus iterative in nature. The polarization energy includes also a self-interaction energy term, the energy required to assemble the set of dipoles

$$E^{\text{efc/efd}} = \sum_{T \neq S}^{\text{Nsol}} Q_T \nabla_S \left(\frac{1}{R_{TS}} \right) \cdot \mathbf{D}_S = \sum_{T \neq S}^{\text{Nsol}} Q_T \frac{(\mathbf{R}_T - \mathbf{R}_S)}{R_{TS}^3} \cdot \mathbf{D}_S, \tag{A53}$$

$$E^{\text{efd/efd}} = \frac{1}{2} \sum_{T \neq S}^{\text{Nsol}} \mathbf{D}_T \cdot \nabla_T \times \nabla_S \left(\frac{1}{R_{TS}} \right) \cdot \mathbf{D}_S$$

$$= \frac{1}{2} \sum_{T \neq S}^{\text{Nsol}} \mathbf{D}_T \cdot \left\{ + \frac{1}{R_{TS}^3} - \frac{3(\mathbf{R}_T - \mathbf{R}_S) \times (\mathbf{R}_T - \mathbf{R}_S)}{R_{TS}^5} \right\} \cdot \mathbf{D}_S, \tag{A54}$$

$$E^{\text{efd/self}} = \frac{1}{2} \sum_T^{\text{Nsol}} \mathbf{D}_T \frac{1}{\alpha_T} \mathbf{D}_T. \tag{A55}$$

where α_T is the magnitude of the isotropic polarizability tensor assigned to center T. In the end we have

$$E_{\text{pol}} = E^{\text{efc/efd}} + E^{\text{efd/efd}} + E^{\text{efd/self}}. \tag{A56}$$

6. The polarization response equations

The polarization response equations yield the induced dipoles at the polarization centers S of the MM subsystem. These centers are assigned a polarizability tensor, isotropic in the present case, and denoted α_S, already seen in Eq. (A55). Assuming linear response, the induced dipole $\mathbf{D}_S$ at center S is related to the polarizability α_S and to the electric field $\mathbf{F}_S$ at center S

$$\mathbf{D}_S = \alpha_S \mathbf{F}_S. \tag{A57}$$

We write the field $\mathbf{F}_S$ as

$$\mathbf{F}_S^{\text{stat}} = \mathbf{F}_S^{\text{atm}} + \mathbf{F}_S^{\text{elec}} + \mathbf{F}_S^{\text{efc}}, \tag{A58}$$

$$\mathbf{F}_S^{\text{dyn}} = \mathbf{F}_S^{\text{efd}}, \tag{A59}$$

$$\mathbf{F}_S = \mathbf{F}_S^{\text{stat}} + \mathbf{F}_S^{\text{dyn}}. \tag{A60}$$

$\mathbf{F}^{\text{atm}}$ is the field from the QM nuclei, $\mathbf{F}^{\text{elec}}$ is the field from the QM electron density, $\mathbf{F}^{\text{efc}}$ is the field from the MM effective charges, and $\mathbf{F}^{\text{efd}}$ is the field from the MM "other" induced dipoles. Expressions for the various terms are

$$\mathbf{F}_S^{\text{atm}} = \sum_A^{\text{Nuc}} -\nabla_S\!\left(\frac{1}{R_{AS}}\right) Z_A = \sum_A^{\text{Nuc}} -\frac{(\mathbf{R}_A - \mathbf{R}_S)}{R_{AS}^3} Z_A, \tag{A61}$$

$$\mathbf{F}^{\text{elec}}(1) = -\nabla_1\!\left(\frac{-1}{r_{1S}}\right) = +\frac{(\mathbf{r}_1 - \mathbf{R}_S)}{\mathbf{r}_{1S}^3}, \tag{A62}$$

$$\mathbf{F}_{\mu\nu}^{\text{elec}} = \int \chi_\mu(1)\, \mathbf{F}^{\text{elec}}(1)\, \chi_\nu(1)\, d\tau_1, \tag{A63}$$

$$\mathbf{F}_S^{\text{elec}} = \sum_{\mu\nu} P_{\mu\nu} \mathbf{F}_{\mu\nu}^{\text{elec}}, \tag{A64}$$

$$\mathbf{F}_S^{\text{efc}} = \sum_T^{\text{Nsol}} -\nabla_S\!\left(\frac{1}{R_{TS}}\right) Q_T = \sum_T^{\text{Nsol}} -\frac{(\mathbf{R}_T - \mathbf{R}_S)}{R_{TS}^3} Q_T, \tag{A65}$$

$$\mathbf{F}_S^{\text{efd}} = \sum_{T \neq S}^{\text{Nsol}} -\nabla_S \times \nabla_T\!\left(\frac{1}{R_{TS}}\right) \cdot \mathbf{D}_T$$
$$= \sum_{S \neq T}^{\text{Nsol}} -\left\{ +\frac{1}{R_{TS}^3} - \frac{3(\mathbf{R}_S - \mathbf{R}_T) \times (\mathbf{R}_S - \mathbf{R}_T)}{R_{TS}^5} \right\} \cdot \mathbf{D}_T. \tag{A66}$$

Atomic integrals of the field operator in Eq. (A63) and the total field in Eq. (A64) can be conveniently evaluated for Gaussian basis functions of any angular momentum by means of a Gaussian transform of the field operator $\nabla(1/r)$ as written in Eq. (A62). See Appendix C.

The equations, given in Eq. (A57), for all the S solvent centers are solved iteratively to self consistency. Initial values of $\mathbf{D}_S$ are obtained by using $F_S = \mathbf{F}_S^{\text{stat}}$. The current values of $\mathbf{D}_S$ are then substituted into Eq. (A66), yielding $\mathbf{F}^{\text{efd}}$ of Eq. (A59) and $\mathbf{F}_S$ of Eq. (A60), with the new values of the field going into Eq. (A57) to give an updated set of induced dipoles $\mathbf{D}_S$.

At this point it is convenient to gather all the energy terms that involve the MM induced dipoles. We denote $E_{\text{ind dip}}$ a quantity made up of these terms. Note that all of them are already accounted for in E_{qm} and in E^{pol}:

$$E_{\text{ind dip}} = E^{\text{atm/efd}} + E^{\text{elec/efd}} + E^{\text{efc/efd}} + E^{\text{efd/efd}} + E^{\text{self/efd}}. \tag{A67}$$

We substitute in Eq. (A67) the expressions for all five terms, as written in Eqs. (A31), (A37), (A40), (A43), (A53), (A54), and (A55). We write the conditions for $E_{\text{ind dip}}$ to be a minimum with respect to variations in the induced dipoles $\mathbf{D}_S$. We obtain a set of equations

$$\left\{ \frac{\partial E_{\text{ind dip}}}{\partial \mathbf{D}_S} = 0 \right\}. \tag{A68}$$

They have the form

$$\left\{ -\mathbf{F}_S^{\text{atm}} - \mathbf{F}_S^{\text{elec}} - \mathbf{F}_S^{\text{efc}} - \mathbf{F}_S^{\text{efd}} + \frac{\mathbf{D}_S}{\alpha_S} = 0 \right\},$$

for all $S = 1, ..., N\text{sol}.$ $\tag{A69}$

These equations are identical to the linear response equations in the form of Eq. (A57). In other words the response equations yield induced dipoles that correspond to a minimum of the total QM/MM energy. This is an important property of the $\mathbf{D}_S$ that greatly simplifies the calculation of the gradients of the total QM/MM energy with respect to the QM Cartesian coordinates and the MM Cartesian Coordinates.[56]

7. The E_{vib} term

Finally we include a vibrational energy term for all MM water molecules. The MM water molecules are not constrained to maintain the experimental water geometry. They are allowed to adopt distorted geometries. The vibrational energy of each molecule is calculated from the quartic force field potential of Bartlett et al.,[62] denoted $E_{\text{vib}}(w)$

$$E_{\text{vib}} = \sum_w^{N\text{wat}} E_{\text{vib}}(w). \tag{A70}$$

In Eq. (A70) w denotes the water solvent molecules, N_{wat} is the number of these water molecules. The separation of $E_{\text{mm/mm}}$ and E_{vib} implies an approximation. The partial charges and the LJ parameters that form the $tip3p$, $pol1$, and $pol2$ potentials,[46,47,61] for example, were derived on the basis of rigid water molecules. We take the same parameters with distorted water molecules. As long as the distortions stay small compared to the intermolecular distances the approximation should be satisfactory. That this is the case is borne out by the optimized structures of hybrid water clusters[49] where bond lengths and bond angle distortions are in qualitative accord with the quantum dimers for example. Earlier, Lie and Clementi[48] used this approach successfully.

8. The $E_{(qm+mm)/cav}$ term

This term is an $ad\ hoc$ repulsive potential whose purpose is to prevent the QM and MM molecules from drifting away during a molecular dynamics simulation. We accomplish this by surrounding the QM/MM system by a "soft" sphere of radius R_{sph}, centered at a given point taken as the origin of the frame. In the spirit of the LJ 6-12 potentials commonly used to represent the repulsive interaction between the QM and MM subsystems as in Eqs. (A45)–(A47) or between the MM molecules as in Eqs. (A50)–(A52) we assign a σ_{cav} parameter and a ε_{cav} parameter to the sphere. In the present work we used the LJ parameters for the oxygen atom in the $tip3p$ potential of water. Using the properties of the LJ expression, the thickness of the spherical wall is given by

$$R_{\text{wall}} = \sigma_{\text{cav}} \times 2^{1/6}. \tag{A71}$$

The radius of the cavity is

$$R_{\text{cav}} = R_{\text{sph}} + R_{\text{wall}}. \tag{A72}$$

We denote I any atom A of the QM subsystem or S of the MM subsystem, R_I its distance to the center of the spherical cavity, and $R_{I\text{cav}}$ its distance to the wall of the cavity. We have

$$R_{I\text{cav}} = R_{\text{cav}} - R_I. \tag{A73}$$

We define some LJ quantities, including the distance to the cavity wall that corresponds to the minimum in the LJ interaction

$$\varepsilon_{I\text{cav}} = (\varepsilon_I \varepsilon_{\text{cav}})^{1/2}, \tag{A74}$$

$$\sigma_{I\text{cav}} = (\sigma_I \sigma_{\text{cav}})^{1/2}, \tag{A75}$$

$$R_{I\text{cav}}^* = \sigma_{I\text{cav}} \times 2^{1/6}. \tag{A76}$$

The interaction energy $E_{(\text{qm}+\text{mm})/\text{cav}}$ is defined as

$$E_{(\text{qm}+\text{mm})/\text{cav}} = \sum_{I}^{\text{Nuc}+N\text{sol}} E_{I\text{cav}}^{\text{vdW}}, \tag{A77}$$

$$E_{I\text{cav}}^{\text{vdW}} = 4\varepsilon_{I\text{cav}} \left[\left(\frac{\sigma_{I\text{cav}}}{R_{I\text{cav}}} \right)^{12} - \left(\frac{\sigma_{I\text{cav}}}{R_{I\text{cav}}} \right)^{6} \right] + \varepsilon_{I\text{cav}},$$

$$\text{if } R_{I\text{cav}} \leqslant R_{I\text{cav}}^*, \tag{A78}$$

$$E_{I\text{cav}}^{\text{vdW}} = 0 \quad , \quad \text{if } R_{I\text{cav}} \geqslant R_{I\text{cav}}^*. \tag{A79}$$

In other words for any atom inside the sphere, if its distance to the hard wall of the cavity is larger than some LJ minimum energy distance, then its interaction with the sphere is set to zero. If this distance is larger than the LJ minimum energy distance, then this atom feels the repulsive part of the LJ potential.

APPENDIX B: QM/MM-pol-vib/CAV GRADIENTS

The starting point for the gradient expressions for $E^{\text{QM/MM}}$ of Eq. (A1) and Eqs. (A44), (A45), (A48), and (A56). $E^{\text{QM/MM}}$ is a function of the Cartesian coordinates of the QM subsystem and those of the MM subsystem, as well as a function of the MM induced dipoles, $E^{\text{QM/MM}} \equiv E^{\text{QM/MM}}(\mathbf{x}_{\text{QM}}, \mathbf{x}_{\text{MM}}, \mathbf{D}_{\text{MM}})$. Let q denote any of the QM or MM atomic coordinates. We have

$$\frac{dE^{\text{QM/MM}}}{dq} = \frac{\partial E^{\text{QM/MM}}}{\partial q} + \frac{\partial E^{\text{QM/MM}}}{\partial \mathbf{D}_{MM}} \times \frac{\partial \mathbf{D}_{\text{MM}}}{\partial q}. \tag{A80}$$

We indicated above that $\partial E^{\text{QM/MM}}/\partial \mathbf{D}_{\text{MM}} \equiv 0$, since the linear response of the induced dipoles minimizes the total energy $E^{\text{QM/MM}}$. Thus it suffices to be concerned only with the explicit dependence of the energy expression with respect to the coordinate q, and we can ignore how the MM induced dipoles depend on q.

Following the well-established procedure for deriving the gradient expression for an HF SCF, or KS SCF, or MC SCF energy as in Eq. (A21) by taking advantage of the self-consistency property of the wave function, we have

$$\frac{\partial E}{\partial q} = \sum_{\mu\nu} \gamma_{\mu\nu} \frac{\partial}{\partial q} (\mu|h|\nu) + \frac{1}{2} \sum_{\mu\nu\lambda\sigma} \Gamma_{\mu\nu\lambda\sigma} \frac{\partial}{\partial q} (\mu\nu|\lambda\sigma)$$

$$- \sum_{\mu\nu} W_{\mu\nu} \frac{\partial}{\partial q} (\mu|\nu), \tag{A81}$$

where $W_{\mu\nu}$ is a Lagrangian-weighted density matrix element

$$W_{\mu\nu} = \sum_{i,j}^{occ} 2\varepsilon_{ij} c_{\mu i} c_{\nu j}. \tag{A82}$$

Where Eq. (A81) differs from a traditional "gas-phase" gradient expression is in the first term on the right-hand side that involves derivatives of the one-electron integrals

$$\frac{\partial}{\partial q}(\mu|h|\nu) = \left(\frac{\partial\mu}{\partial q}|h|\nu \right) + \left(\mu|\frac{\partial h}{\partial q}|\nu \right) + \left(\mu|h|\frac{\partial\nu}{\partial q} \right). \tag{A83}$$

When q is an MM coordinate the first term and the last term in Eq. (A83) are identically zero. The four operators given in Eqs. (A27)–(A30) contribute to $\partial/\partial q (\mu|h|\nu)$. V^{atm} of Eqs. (A28), (A34) and V^{efc} of Eqs. (A29) and (A35) involve the same type of integrals. Note that V^{atm} contributes to the gradient for the QM susbsystem only, while V^{efc} contributes to the MM gradient as well as to the QM gradient through the derivatives of the basis functions μ and ν. V^{efd} of Eqs. (A30) and (A36) involve somewhat more complex integrals and derivatives, all of which can be handled in a systematic fashion through the use of the generalized Rys quadrature approach outlined in Appendix C and presented in detail in Ref. (A43). V^{efd} contributes also to the QM gradient and to the MM gradient.

The derivatives of $E^{\text{atm/atm}}$, $E^{\text{atm/efc}}$, and $E^{\text{atm/efd}}$ that come into play as a result of Eq. (A44), are straightforwardly evaluated. Similarly for $E_{\text{qm/mm}} \equiv E_{\text{qm/mm}}^{\text{vdW}}$ and for the various terms that form $E_{\text{mm/mm}}$ according to Eqs. (48) and (56). The derivatives of the vibrational energy term E_{vib} of Eq. (70) and of $E_{(\text{qm}+\text{mm})/\text{cav}}$ are also straightforward to handle.

It is good to remember here again that a great simplification has come into the gradient derivation owing to the fact that the MM induced dipoles $\mathbf{D}_S$ minimize the energy of the total QM/MM system.[56] As a result there was not need to be concerned with the explicit differentiation of the energy with respect to the induced dipoles.

APPENDIX C: GENERALIZED Rys QUADRATURE FOR FIELD AND FIELD GRADIENT INTEGRALS AND THEIR DERIVATIVES

Transforms for field and field gradient integrals can be obtained from the general expression

$$\int_0^\infty u^{2n} e^{-ru^2} du = \frac{1 \cdot 3 \cdots (2n-1)}{2^{n+1}} \left(\frac{\pi}{r^{2n+1}} \right)^{1/2}. \tag{A84}$$

Alternately it suffices to differentiate to the order needed, the Gaussian transform expression for $1/r_{1C}$ with respect to the probe center. It follows that

$$\frac{1}{r_{1S}} = \frac{2}{\sqrt{\pi}} \int_0^\infty e^{-r_{1S}^2 u^2} du, \tag{A85}$$

$$\frac{x_{1S}}{r_{1S}^3} = \frac{2}{\sqrt{\pi}} \int_0^\infty 2x_{1S} u^2 e^{-r_{1S}^2 u^2} du, \tag{A86}$$

$$-\frac{1}{r_{1S}^3} + 3\frac{x_{1S}x_{1S}}{r_{1S}^5} = \frac{2}{\sqrt{\pi}} \int_0^\infty (-2u^2 + 4x_{1S}^2 u^4) e^{-r_{1S}^2 u^2} du, \tag{A87}$$

$$3\frac{x_{1S}y_{1S}}{r_{1S}^5} = \frac{2}{\sqrt{\pi}} \int_0^\infty 4x_{1S}y_{1S}u^4 e^{-r_{1S}^2 u^2} du. \tag{A88}$$

Subroutines for the calculations of these and other operators could be found in programs such as HONDO[57] and GAMESS[68] for some time. Systematic extensions for the calculation of their derivatives with respect to center S are given by Dupuis.[67]

[1] J. Tomasi and M. Persico, Chem. Rev. **94**, 2027 (1994).

[2] C. J. Cramer and D. G. Truhlar, Chem. Rev. **99**, 2161 (1999).

[3] J. Gao, Acc. Chem. Res. **29**, 298 (1996).

[4] J. Gao, Rev. Comput. Chem. **7**, 119 (1996).

[5] D. Rinaldi and J. L. Rivail, Theor. Chim. Acta **32**, 57 (1973).

[6] M. Karelson, A. R. Katritzki, M. Szafran, and M. C. Zerner, J. Org. Chem. **54**, 6030 (1989).

[7] C. Amovilli, V. Barone, R. Cammi, E. Cances, M. Cossi, B. Mennucci, C. S. Pomelli, and J. Tomasi, Angew. Chem. **32**, 227 (1999).

[8] A. Klamt and G. Schuurmann, J. Chem. Soc., Perkin Trans. 2 **2**, 799 (1993).

[9] M. Cossi, B. Mennucci, and R. Cammi, J. Comput. Chem. **17**, 57 (1996).

[10] J. B. Foresman, T. A. Keith, K. B. Wiberg, J. Snoonian, and M. J. Frisch, J. Phys. Chem. **100**, 16098 (1996).

[11] C. G. Zhan and D. N. Chipman, J. Chem. Phys. **109**, 10543 (1998).

[12] D. N. Chipman and M. Dupuis, Theor. Chem. Acc. **107**, 90 (2002).

[13] V. Barone, M. Cossi, and J. Tomasi, J. Chem. Phys. **107**, 3210 (1997).

[14] C. Cappelli, B. Mennucci, C. O. da Silva, and J. Tomasi, J. Chem. Phys. **112**, 5382 (2000).

[15] M. A. Aguilar, F. J. Olivares del Valle, and J. Tomasi, J. Chem. Phys. **98**, 7375 (1993).

[16] M. Cossi and V. Barone, J. Chem. Phys. **112**, 2427 (2000).

[17] M. A. Aguilar, J. Phys. Chem. A **105**, 10393 (2001).

[18] M. Cossi and V. Barone, J. Chem. Phys. **112**, 2427 (2000).

[19] T. N. Truong, T. T. Truong, and E. Stefanovich, J. Chem. Phys. **107**, 1881 (1997).

[20] T. Truong, J. Phys. Chem. B **102**, 7877 (1998).

[21] M. Cossi, C. Adamo, and V. Barone, Chem. Phys. Lett. **297**, 1 (1998).

[22] A. Warshel and M. Levitt, J. Mol. Biol. **103**, 227 (1976).

[23] M. J. Field, P. A. Bash, and M. Karplus, J. Comput. Chem. **11**, 700 (1990).

[24] J. Gao and X. Xia, Science **258**, 631 (1992).

[25] J. Gao, J. Phys. Chem. **96**, 537 (1992).

[26] P. Bash, M. J. Field, and M. Karplus, J. Am. Chem. Soc. **109**, 8092 (1987).

[27] A. H. de Vries, P. Th. Van Duijnen, A. H. Huffer, J. A. C. Rullmann, J. P. Dijkman, H. Merenga, and B. T. Thole, J. Comput. Chem. **16**, 37 (1995); P. Th. van Duijnen, F. Grozema, and M. Swart, J. Mol. Struct.: THEOCHEM **464**, 191 (1999).

[28] R. V. Stanton, D. V. Hartsough, and K. N. Merz, Jr., J. Phys. Chem. **97**, 11868 (1993).

[29] I. Tunon, M. T. C. Martins-Costa, C. Millot, M. F. Ruiz-Lopez, and J. L. Rivail, J. Comput. Chem. **17**, 19 (1996).

[30] M. L. Sanchez, M. A. Aguilar, and F. J. Olivares del Valle, J. Comput. Chem. **18**, 313 (1997).

[31] S. Ten-no, F. Hirata, and S. Kato, Chem. Phys. Lett. **214**, 391 (1993).

[32] S. Ten-no, F. Hirata, and S. Kato, J. Chem. Phys. **100**, 7443 (1994).

[33] M. Freindorf and J. Gao, J. Comput. Chem. **17**, 386 (1996).

[34] T. N. Truong and E. V. Stefanovich, Chem. Phys. Lett. **256**, 348 (1996).

[35] T. N. Truong and E. V. Stefanovich, Chem. Phys. **218**, 231 (1997).

[36] M. A. Thompson and G. K. Schenter, J. Phys. Chem. **99**, 6374 (1995).

[37] M. A. Thompson, J. Phys. Chem. **100**, 14492 (1996).

[38] J. Gao, J. Comput. Chem. **18**, 1061 (1997).

[39] R. A. Bryce, R. Buesnel, I. H. Hillier, and N. A. Burton, Chem. Phys. Lett. **279**, 367 (1997).

[40] M. S. Gordon, M. A. Freitag, P. Bandyopadhyay, J. H. Jensen, V. Kairys, W. J. Stevens, J. Phys. Chem. A **105**, 293 (2001).

[41] A. Smith, M. A. Vincent, and I. H. Hillier, J. Phys. Chem. A **103**, 1132 (1999).

[42] B. L. Trout and M. Parrinello, Chem. Phys. Lett. **288**, 343 (1998).

[43] B. L. Trout and M. Parrinello, J. Phys. Chem. B **103**, 7340 (1999).

[44] P. L. Geissler, C. Dellago, D. Chandler, J. Hutter, and M. Parrinello, Science **291**, 2121 (2001).

[45] R. Car and M. Parrinello, Phys. Rev. Lett. **55**, 2471 (1985).

[46] L. X. Dang, J. Chem. Phys. **97**, 2659 (1992).

[47] L. X. Dang and T. M. Chang, J. Chem. Phys. **106**, 8149 (1997).

[48] G. C. Lie and E. Clementi, Phys. Rev. A **33**, 2679 (1986).

[49] M. Aida, H. Yamataka, and M. Dupuis, Int. J. Quantum Chem. **77**, 199 (2000).

[50] M. Aida, H. Yamataka, and M. Dupuis, Chem. Phys. Lett. **292**, 474 (1998).

[51] M. Aida, H. Yamataka, and M. Dupuis, Theor. Chem. Acc. **102**, 262 (1999).

[52] R. M. Levy, D. B. Kitchen, J. T. Blair, and K. Krogh-Jespersen, J. Phys. Chem. **94**, 4470 (1990).

[53] M. Dupuis, Y. Kawashima, and K. Hirao, J. Chem. Phys. **117**, 1256 (2002) (next paper in this issue).

[54] Y. Kawashima, K. Hirao, and M. Dupuis (in preparation).

[55] J. M. Bader, C. M. Cortis, and B. J. Berne, J. Chem. Phys. **106**, 2372 (1997).

[56] J. G. Angyan, F. Colonna-Cesari, and O. Tapia, Chem. Phys. Lett. **166**, 180 (1990).

[57] M. Dupuis, A. Marquez, and E. R. Davidson, HONDO-2001, based on HONDO-95. available from QCPE, Indiana University.

[58] S. A. Maluendes and M. Dupuis, Int. J. Quantum Chem. **42**, 1327 (1992).

[59] W. J. Hehre, R. Ditchfield, and J. A. Pople, J. Chem. Phys. **56**, 2257 (1972); P. C. Hariharan and J. A. Pople, Theor. Chim. Acta **28**, 213 (1973).

[60] T. H. Dunning, Jr., J. Chem. Phys. **90**, 1007 (1989); **90**, 1007 (1989).

[61] W. L. Jorgensen, J. Am. Chem. Soc. **103**, 335 (1981).

[62] R. J. Bartlett, I. Shavitt, and G. D. Purvis, J. Chem. Phys. **71**, 281 (1979).

[63] B. O. Roos, P. R. Taylor, and P. E. M. Siegbahn, Chem. Phys. **48**, 157 (1980); the CAS implementation in the HONDO code is described in M. Dupuis, in *Modern Techniques in Computational Chemistry: MOTECC-91*, edited by E. Clemeti (ESCOM Science Publishers B. V., The Netherlands 1991), p. 469.

[64] S. F. Boys and F. Bernardi, Mol. Phys. **19**, 553 (1970).

[65] M. Dupuis (in preparation).

[66] J. A. Pople, P. M. W. Gill, and B. G. Johnson, Chem. Phys. Lett. **199**, 557 (1992).

[67] M. Dupuis, Comput. Phys. Commun. **134**, 150 (2001).

[68] M. W. Schmidt, K. K. Baldridge, and J. A. Boatz *et al.*, J. Comput. Chem. **14**, 1347 (1993).