The merits of the frozen-density embedding scheme to model solvatochromic shifts

Johannes Neugebauer,[a] Manuel J. Louwerse,[b] and Evert Jan Baerends[c]
Theoretical Chemistry, Vrije Universiteit Amsterdam, De Boelelaan 1083, 1081 HV Amsterdam, The Netherlands

Tomasz A. Wesolowski[d]
Department of Physical Chemistry, University of Geneva, 30 quai Ernest-Ansermet, CH-1211 Geneva 4, Switzerland

(Received 8 December 2004; accepted 21 December 2004; published online 3 March 2005)

We investigate the usefulness of a frozen-density embedding scheme within density-functional theory [J. Phys. Chem. **97**, 8050 (1993)] for the calculation of solvatochromic shifts. The frozen-density calculations, particularly of excitation energies have two clear advantages over the standard supermolecule calculations: (i) calculations for much larger systems are feasible, since the time-consuming time-dependent density functional theory (TDDFT) part is carried out in a limited molecular orbital space, while the effect of the surroundings is still included at a quantum mechanical level. This allows a large number of solvent molecules to be included and thus affords both specific and nonspecific solvent effects to be modeled. (ii) Only excitations of the system of interest, i.e., the selected embedded system, are calculated. This allows an easy analysis and interpretation of the results. In TDDFT calculations, it avoids unphysical results introduced by spurious mixings with the artificially too low charge-transfer excitations which are an artifact of the adiabatic local-density approximation or generalized gradient approximation exchange-correlation kernels currently used. The performance of the frozen-density embedding method is tested for the well-studied solvatochromic properties of the $n \rightarrow \pi^*$ excitation of acetone. Further enhancement of the efficiency is studied by constructing approximate solvent densities, e.g., from a superposition of densities of individual solvent molecules. This is demonstrated for systems with up to 802 atoms. To obtain a realistic modeling of the absorption spectra of solvated molecules, including the effect of the solvent motions, we combine the embedding scheme with classical molecular dynamics (MD) and Car-Parrinello MD simulations to obtain snapshots of the solute and its solvent environment, for which then excitation energies are calculated. The frozen-density embedding yields estimated solvent shifts in the range of 0.20–0.26 eV, in good agreement with experimental values of between 0.19 and 0.21 eV. © *2005 American Institute of Physics*. [DOI: 10.1063/1.1858411]

I. INTRODUCTION

Solvatochromism denotes the shift of absorption or emission bands of solutes in solvents of different polarity.[1–3] These shifts are caused by a different stabilization of ground and excited states by the solvent within the Franck–Condon region, i.e., without a reorientation of the solvent molecules upon absorption or emission. In recent time, several attempts have been made to model the solvatochromic shifts of simple compounds such as acetone,[4,5] triazines or tetrazines,[6–8] solvatochromic organic dyes,[9–11] or transition metal compounds[12–15] with different explicit or implicit models for the solvent.

While implicit models, i.e., continuum solvation models, are quite successful to describe nonspecific effects, such as dielectric medium effects,[16–18] explicit models are needed for describing specific solvent effects, e.g., hydrogen bonding. An efficient example for explicit models are quantum mechanics/molecular mechanics (QM/MM) Car–Parrinello molecular dynamics (CPMD) schemes[19] in combination with TDDFT or restricted open-shell Kohn–Sham calculations as presented in Refs. 5 and 11. A fully quantum mechanical treatment of the solvent is, in principle, possible using CPMD, with a subsequent TDDFT calculation of excitation energies for snapshots of the simulation, as has been demonstrated for *s*-tetrazine[6] or acetone.[20,21] In these calculations, however, the problem of spurious or artificially too low charge-transfer (CT) excitations[22–24] at low energies within TDDFT makes it cumbersome to determine the actual excitations of interest[6] although at least the problem of solute → solvent excitations could be partially removed by including Hartree–Fock exchange in the exchange-correlation potential.[21,25] This problem is related to the fact that the difference between the true excitation energy for a CT excitation over a long distance and the corresponding TDDFT result is [in the adiabatic local-density approximation (LDA) approximation] essentially[24]

[a]Electronic mail: jneugeb@chem.vu.nl
[b]Electronic mail: louwerse@chem.vu.nl
[c]Electronic mail: baerends@chem.vu.nl
[d]Electronic mail: Tomasz.Wesolowski@chiphy.unige.ch

$$\Delta^A = \epsilon_A^- - \epsilon_A, \tag{1}$$

where ϵ_A^- is the energy of the highest occupied molecular orbital (HOMO) of the anionic acceptor fragment, while ϵ_A is the energy of the lowest unoccupied molecular orbital (LUMO) of the neutral acceptor fragment (cf. Ref. 24). As will be explicitly demonstrated in this study, the problem becomes the more severe, the more solvent molecules are involved, making the supermolecule calculations practically impossible. Even in those cases, where supermolecule calculations are still possible, it may be very difficult to extract the properties of the solvated molecule.[26]

The orbital-free embedding formalism, in which the electron density of the embedded system is obtained from Kohn–Sham-like one-electron equations in the field of a "frozen" environment density (Kohn–Sham approach with constrained electron density, KSCED),[27] and its time-dependent response generalization[28] can be of great value in this context: No (or only a few) solvent molecules are explicitly considered in terms of their orbitals, so that no (or only a few) spuriously low excitations occur. Still, the full effect of the solvent molecules' density on the potential is included. Therefore, the KSCED embedding offers the possibility to describe both specific solvent effects, since the structure of the solvent is explicitly modeled, and general, non-specific solvent effects since its efficiency allows to include a large number of solvent molecules. Its ability to model hydrogen-bonding induced shifts of excitation energies has recently been demonstrated for nucleic acid base pairs.[29] An alternative method is used in Refs. 30 and 31, which explicitly takes the full system into account during the self-consistent-field results (SCF) steps, but restricts the orbital space in the TDDFT step to those orbitals important for the solute. Our method uses the frozen-density embedding during the SCF procedure, enabling very large environments to be treated, see below.

The aim of this study is to elucidate the advantages and disadvantages of replacing the explicit treatment of solvent molecules in supermolecule calculations by the frozen-density approach in the KSCED embedding, and to investigate its ability to efficiently model solvatochromic shifts by performing TDDFT calculations for systems with many solvent molecules for many configurations. As an example, we choose the solvent shift of the $n \rightarrow \pi^*$ excitation of acetone in water, since it represents a benchmark for methods to model solvatochromic shifts. Many experimental[32–36] and theoretical studies[4,5,21] for this system have been carried out so that reliable reference data are available. This solvent shift has been calculated in Ref. 4 using dielectric continuum models and explicit, polarizable molecular solvent models in a QM/MM scheme. In that work, a solvent shift of $+0.23 \pm 0.04$ eV in water was calculated using the latter, while the continuum models led to much worse results. The experimental shift is between 0.19 and 0.21 eV.[32–36] A recent QM/MM simulation[5] using the restricted open-shell Kohn–Sham approach for the QM part led to a solvent shift of +0.25 eV in water. An entirely quantum mechanical CPMD simulation[21] with a small periodic box of one acetone molecule and 14 water molecules resulted in an approximate solvent shift of 0.206 eV.

This work is organized as follows. In Sec. II a brief introduction into the theoretical background of frozen-density KSCED calculations in combination with linear response TDDFT calculations is given. Methods for further approximations for an efficient calculation of solvent electron densities are studied in Sec. III. A comparison of KSCED embedding excitation energies with supermolecule calculations, which elucidates the problem of identifying in the supermolecule calculations valence excitations among the CT excitations from water lone-pair orbitals to the acetone π^* is carried out in Sec. IV. In Sec. V finally, we use statistical averaging over the solvent molecule motions, applying the frozen-density embedding on snapshots from (CP)MD simulations, in order to obtain a complete simulation of the $n \rightarrow \pi^*$ solvent shift of acetone in water.

II. METHODOLOGY

In the orbital-free embedding formalism[27] the electron density of the embedded subsystem (ρ_I) in a given microscopic environment, which is represented by means of the frozen electron density (ρ_{II}) and a set of nuclear charges ($Z_{A_{II}}$) at the corresponding positions ($\mathbf{R}_{A_{II}}$), is derived from Kohn–Sham-like one-electron equations. The effective potential in these equations is derived from the requirement that the total density $\rho_{total} = \rho_I + \rho_{II}$ of the system is obtained, which minimizes the total energy, from an optimization process in which the environment density ρ_{II} is kept frozen. On the assumption that the complementary ρ_I is positive definite, and is noninteracting v_s representable, one can derive KS-type equations in which the effect of ρ_{II} is folded into the effective potential for the ρ_I system. We refer to this effective potential and the corresponding equations as KSCED. The KSCED effective potential contains the same terms as the isolated system I would—kinetic energy operator and nuclear attraction, electron Coulomb, and exchange-correlation potentials—plus an additional (embedding) component which reads

$$\begin{aligned} V_{\mathrm{emb}}^{\mathrm{eff}}[\mathbf{r}, \rho_I, \rho_{II}] = \sum_{A_{II}} & -\frac{Z_{A_{II}}}{|\mathbf{r} - \mathbf{R}_{A_{II}}|} + \int \frac{\rho_{II}(\mathbf{r}')}{|\mathbf{r}' - \mathbf{r}|} d\mathbf{r}' \\ & + \left. \frac{\delta E_{xc}[\rho]}{\delta \rho} \right|_{\rho = \rho_I + \rho_{II}} - \left. \frac{\delta E_{xc}[\rho]}{\delta \rho} \right|_{\rho = \rho_I} \\ & + \left. \frac{\delta T_s[\rho]}{\delta \rho} \right|_{\rho = \rho_I + \rho_{II}} - \left. \frac{\delta T_s[\rho]}{\delta \rho} \right|_{\rho = \rho_I}, \end{aligned} \tag{2}$$

where we may also write the kinetic part of the potential as the functional derivative $\delta T_s^{\mathrm{nad}}[\rho_I, \rho_{II}] / \delta \rho_I$ of the nonadditive kinetic energy functional,

$$T_s^{\mathrm{nad}}[\rho_I, \rho_{II}] = T_s[\rho_I + \rho_{II}] - T_s[\rho_I] - T_s[\rho_{II}]. \tag{3}$$

The functionals $E_{xc}[\rho]$, and $T_s[\rho]$ are defined in the Kohn–Sham formulation of DFT. The above system-independent form of the effective embedding potential was introduced using the subsystem formulation of density-functional theory

introduced originally by Cortona.[37] The exchange-correlation component of the effective potential is approximated using the PW91 exchange-correlation functional. The simplest approximation for the kinetic energy functional, corresponding to the local-density approximation in KS-DFT, is the Thomas–Fermi functional,

$$T_s^{\mathrm{TF}} = C_{\mathrm{TF}} \int \rho^{5/3} d\mathbf{r}. \tag{4}$$

The choice of the kinetic energy functional is of great importance for the performance of the model. An analysis of different approximations for this functional is given in Refs. 38–40. For the construction of these functionals, the suggestion of Lee, Lee, and Parr[41] to use similar analytical forms for kinetic energy and exchange energy functionals is applied there,

$$E_{\mathrm{x}}[\rho] \approx E_{\mathrm{x}}^{\mathrm{GGA}} = \int f(\rho(\mathbf{r}), |\nabla \rho(\mathbf{r})|) d\mathbf{r} \tag{5}$$

$$= - C_{\mathrm{x}} \int \rho^{4/3}(\mathbf{r}) F(s(\mathbf{r})) d\mathbf{r}, \tag{6}$$

and

$$T_s^{\mathrm{GGA}}[\rho] = C_{\mathrm{F}} \int \rho^{5/3}(\mathbf{r}) F(s(\mathbf{r})) d\mathbf{r}, \tag{7}$$

where $C_{\mathrm{x}} = (3/4)(3/\pi)^{1/3}$, $C_F = (3/10)(3\pi^2)^{(2/3)}$, and $s = |\nabla \rho|/(2\rho k_{\mathrm{F}})$, with $k_{\mathrm{F}} = (3\pi^2\rho)^{1/3}$.

Studies[39,42] of the accuracy of various approximations to $T_s^{\mathrm{nad}}[\rho_I, \rho_{II}]$ in the case of weakly overlapping pairs of electron densities ρ_I, ρ_{II} showed that the most accurate nonadditive kinetic energy functional (and the associated functional derivative) has the same analytic form of the enhancement factor $F(s)$ as the exchange functional of Perdew and Wang (see Ref. 43) but it is reparametrized for the kinetic energy by Lembarki and Chermette.[44] Its complete form reads

$$F_{\mathrm{LC94}}(s) = \frac{1 + 0.093\,907 s \operatorname{arcsinh}(76.32 s) + (0.266\,08 - 0.080\,9615 e^{-100 s^2}) s^2}{1 + 0.093\,907 s \operatorname{arcsinh}(76.32 s) + 0.577\,67 \times 10^{-4} s^4}. \tag{8}$$

Equations (3), (7), and (8) lead to the nonadditive kinetic energy functional used also in this work, dubbed PW91k.

The time-dependent response formulation of the KSCED scheme is given in Ref. 28. Using the orbitals and eigenvalues from the KSCED equations, the only additional requirement for the TDDFT-KSCED calculations is the adaptation of the exchange-correlation kernel (for system *A*), which now also contains a Thomas–Fermi component.

We will compare below the TDDFT-KSCED calculations of the excitation energies of a solute molecule employing the frozen-density representation of the solvent environment to supermolecule calculations. The supermolecule results, however, do not provide a rigorous benchmark for the KSCED calculations for the following reasons. (i) The nonadditive part of the kinetic energy is represented by some density functional for the kinetic energy, while in supermolecule calculations, as in all regular KS calculations, such a functional is not needed. So there are essential differences in the density-functional treatment of the kinetic energy in the supermolecule and the KSCED calculation, precluding exact comparison. (ii) We did not include basis functions on system II atoms in our calculations for the embedded molecule, i.e., its density is localized in the embedded part only. The effect may be slight, but exact equality of supermolecule and embedded calculations would require complete basis set flexibility in the ρ_I calculations also in the ρ_{II} region. (iii) Differences also arise from our use in most cases of different exchange-correlation potentials for the frozen and the embedded density. The frozen density can be constructed in various ways, where computational efficiency is an important consideration. The TDDFT calculations on the embedded molecule, however, always employ our most advanced KS potential, the "statistical averaging of (model) orbital potentials" (SAOP) potential, for most accurate excitation energies. We have, thus, chosen a pragmatic approach, which is most appropriate for the system under study here; the used embedding potential was shown to be the most adequate for the case of two subsystems for which the overlap between the corresponding densities is weak. The fact that the supermolecule calculations are not strictly a benchmark for the embedded calculations should be kept in mind, see also earlier work.[38,39,42,45]

All density-functional calculations (apart from the CPMD calculations) have been performed using the Amsterdam density functional (ADF) package.[46,47] We employ the Becke–Perdew functional, dubbed BP86,[48,49] in combination with the TZP basis set, a triple-ζ basis with one set of polarization functions, from the ADF basis set library[46] for structure optimizations. The SAOP potential[50–52] is used with the TZP basis set to calculate vertical TDDFT excitation energies, since it is well suited for describing both valence and Rydberg excited states; this is in contrast to normal generalized gradient approximation (GGA) potentials, which introduce additional problems by artificially placing Rydberg excitations at much too low energies.[53] Comparisons with calculations using the larger QZ3P-1DIFFUSE basis set, which is of quadruple-ζ quality and contains three sets of polarization functions as well as additional diffuse functions, showed that the excitation energies of the transitions investigated in this study change by only 0.03–0.06 eV compared to the TZP basis.

For classical MD simulations, we employed the general Amber force field (GAFF)[54] and the TIP3P water model[55] using the TINKER package[56,57] to run the calculations. The

CPMD simulations were performed with the projector augmented wave (PAW) package[58] using the BLYP exchange-correlation functional.[48,59]

III. CONSTRUCTION OF THE FROZEN ENVIRONMENT DENSITY

Since a realistic modeling of solvent effects requires that a large number of surrounding water molecules is taken into account, we wish to perform the calculation of the frozen environment density as simply as possible. For this purpose, we tested different construction methods for the solvent density for two different systems. The first test system is an energy-minimized cluster of acetone with two water molecules, while the second is a much larger cluster of acetone with 52 water molecules, which was obtained as an arbitrary snapshot from a classical MD simulation as explained in Sec. V. Note that here and in the following section, all water molecules are included in the *frozen environment* system, if not explicitly stated otherwise, in order to establish well-defined conditions for the calculation.

For both systems we employed different constructions of the solvent electron density. A fully consistent calculation would require to use the SAOP potential, which we use to calculate the vertical excitation energies of the embedded molecule, also for the construction of the solvent density. For comparison, we also tested the much simpler local-density approximation in this preparation step. As a further parameter, we tested different values for the SCF convergence parameter s_{conv} in the construction of the frozen density. In our study, convergence is considered reached if the maximum element of the commutator of the Fock matrix with the density matrix used to construct this Fock matrix falls below s_{conv} and the norm of the matrix below $10s_{\mathrm{conv}}$. Besides the default value of $s_{\mathrm{conv}}=1\mathrm{E}-6$, we also tested larger values of up to $s_{\mathrm{conv}}=0.5$ for both systems. We note that an "exact" KSCED calculation, including for instance basis functions in the full environment system to describe any change of ρ_{II} towards ρ_{total} in this region, ought to correct automatically for differences in the frozen ρ_{II} density. Since our KSCED calculations are not exact, the different construction methods imply slightly different approximations, the effect of which we are testing in Table I.

Excitation energies are given in Table I for two valence excitations of acetone, the experimentally known $n\rightarrow\pi^*$ excitation and a $\sigma\rightarrow\pi^*$ excitation with very low oscillator strength, which is not known from experiment. The latter is included because it shows a large shift in the calculations, and is therefore instructive for the comparison of Kohn–Sham supermolecule and the frozen-density embedding calculations in Sec. IV.

The results of Table I show that differences between SAOP and LDA densities for the environment are almost negligible: They are between 0.003 and 0.005 eV for a given SCF convergence parameter. Furthermore, it can be seen that even the lowest tested convergence parameter still leads to acceptable results in the KSCED-TDDFT calculation. For the determination of the solvent shifts of acetone presented in Sec. V, and the study of the effect of the number of water molecules on the spectrum in Sec. IV we chose a parameter of $s_{\mathrm{conv}}=0.1$, which resulted in deviations not larger than 0.007 eV for all our tests. This is much smaller than the expected solvent shift. The calculated oscillator strengths are in similarly good agreement.

TABLE I. Excitation energies (SAOP/TZP; in units of eV) of the $n\rightarrow\pi^*$ and $\sigma\rightarrow\pi^*$ valence excitations of acetone in water using the QM/QM embedding scheme for clusters with two water molecules (optimized structure) or 52 water molecules (structure from arbitrary snapshot of a MD simulation). (N.B. Due to the different structures, the 2 H_2O and 52 H_2O results are not comparable.) For the preparation of the frozen densities, either SAOP/TZP and LDA/TZP (two water molecules) or SAOP/DZ and LDA/DZ (52 water molecules) were applied in combination with different SCF convergence parameters s_{conv} (see text for explanation). Additionally, approximate densities from superpositions of molecular fragments ("mol. frags."), either taking the sum of the fragment densities ("sumf.") or the density obtained after one diagonalization of the Fock matrix based on these fragment densities ("diag.") are employed.

		2 H_2O		52 H_2O	
	s_{conv}	SAOP	LDA	SAOP	LDA
$n\rightarrow\pi^*$	$1.0E-6$	4.7841	4.7873	4.8263	4.8215
	$1.0E-3$	4.7841	4.7873	4.8263	4.8215
	$1.0E-1$	4.7841	4.7870	4.8274	4.8222
	$2.0E-1$	4.7841	4.7878	4.8274	4.8222
	$5.0E-1$	4.7739	4.7705	4.8300	4.8227
	Mol. frags. (sumf.)	⋯	4.7943	⋯	4.8171
	Mol. frags. (diag.)	⋯	4.7837	⋯	4.8249
$\sigma\rightarrow\pi^*$	$1.0E-6$	7.2819	7.2906	7.2118	7.2196
	$1.0E-3$	7.2819	7.2907	7.2118	7.2196
	$1.0E-1$	7.2815	7.2909	7.2082	7.2174
	$2.0E-1$	7.2815	7.2904	7.2082	7.2174
	$5.0E-1$	7.2948	7.3157	7.2029	7.2182
	Mol. frags. (sumf.)	⋯	7.2822	⋯	7.2363
	Mol. frags. (diag.)	⋯	7.2944	⋯	7.2141

In another test ("mol. frags." in Table I) we used the density obtained from a superposition of densities of individual solvent molecules as a guess for the density of the full system of solvent molecules. First, for each solvent molecule (which all may have different O–H bond lengths and H–O–H angles) an SCF calculation is performed to obtain the density. Taking just the sum of these fragment densities as frozen environment density is labeled "sumf." in Table I. We also used this superposition to construct the Coulomb and exchange-correlation potentials of the KS operator, constructed a Fock matrix for the full environment system with these operators, and performed one diagonalization to obtain approximate orbitals and density of the full frozen system ("diag." in Table I). The latter is the default for all following discussions. Note that these calculations are not carried out with the SAOP potential, since this potential always requires an intermediate step to construct the orbitals and the density of the system before setting up the potential and the Fock matrix.

The results employing molecular fragment densities are very encouraging with errors smaller than 0.006 eV for the density after one Fock matrix diagonalization, and slightly higher errors for the simple sum of fragment densities (about 0.008 eV for the $n\rightarrow\pi^*$ transition, and about twice as much

for the $\sigma \rightarrow \pi^*$ transition). This is particularly important since the bottleneck in the calculation of excitation energies within the KSCED approach for very large frozen systems is the SCF to prepare the frozen density. This shows the usual N^3 DFT scaling behavior with the system size N, unless the environment system becomes so large that one enters the linear scaling regime.[60] The use of molecular fragments results in a scaling behavior proportional to N with a very small prefactor in the preparation of the density, since one small SCF calculation for each solvent molecule is required (plus the superposition step). Using a "rigid molecule" approach (see Sec. IV C), in which the density is calculated only for a limited number of different solvent molecule structures, of which copies are then used to construct the density of outer solvation shells, the preparation of the frozen density requires a constant computational effort, irrespectively of the number of solvent molecules (not considering the superposition step).

IV. CONVERGENCE WITH THE SIZE OF THE H_2O CLUSTER

The CPMD method will be used in Sec. V to model the solvent shift by a statistics using many solvent configurations from a CPMD run. For the TDDFT calculations using these snapshot configurations within the QM/QM embedding scheme we have to use a cluster ansatz, since our TDDFT implementation does not allow to use periodic boundary conditions.

In order to investigate the convergence of the excitation energies of acetone with respect to the size of the solvation shell, we calculated the lowest excitations for acetone and acetone embedded in clusters of water molecules of increasing size. All structures employed here are substructures of a particular snapshot from a CPMD simulation as explained in Sec. V. Calculations were performed for the isolated acetone molecule in the structure from the snapshot, and for acetone in structures with 2, 5, 10, 20, 57, and 88 water molecules. For the largest system with 88 water molecules (274 atoms, including acetone; shown in Fig. 1), we used a simplified density obtained as a superposition of molecular densities as explained in Sec. III (scheme "diag."). In order to estimate the effect of this approximation for the current snapshot, we used this simplified method also in an additional calculation on the cluster with 57 water molecules, and similar to the observations in Sec. III, this hardly affects the results.

For each of these clusters, excitation energies were calculated in three different ways: (A) A TDDFT supermolecule calculation for the whole system was performed, (B) a TDDFT calculation for the acetone molecule and the two nearest water molecules, which are bound via hydrogen bonds to the carbonyl oxygen, is carried out, whereas the effect of the additional water molecules is included via the KSCED frozen-density embedding, and (C) a TDDFT calculation for the acetone molecule only is performed in which the effect of all water molecules is included via the KSCED frozen-density embedding. For the largest clusters with 57 or 88 water molecules, we only present calculations of type B and C, since supermolecule TDDFT calculations for such large systems are hardly feasible, and more important, the long-range CT problem of TDDFT would introduce an enormously large number of artificially low excitations as will be shown below. We used the LDA and a double-ζ (DZ) basis set as implemented in ADF (Ref. 46) to obtain the frozen density of the solvent molecules. For better comparability, the DZ basis set was also used for solvent molecules in schemes A and B. Results are given for the $n \rightarrow \pi^*$ and $\sigma \rightarrow \pi^*$ excitations of acetone. All excitation energies are presented in Table II, the corresponding shifts with respect to

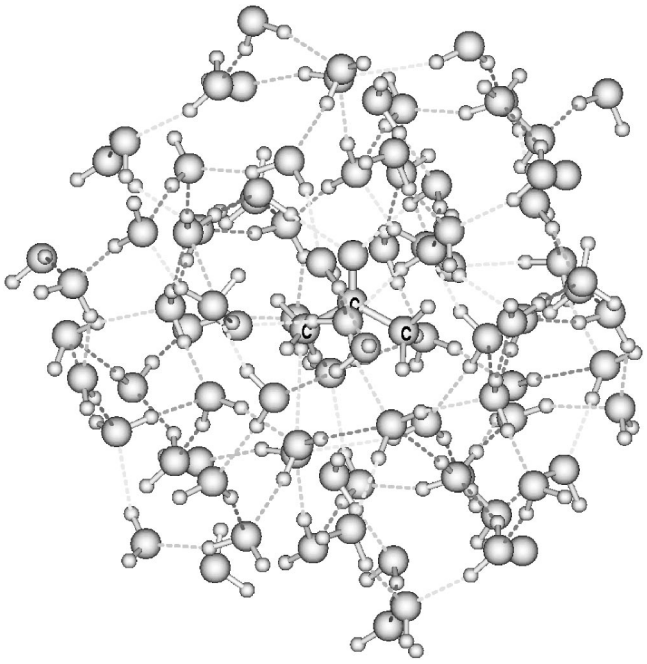

FIG. 1. Example structure of acetone and its 88 nearest water molecules as obtained from a CPMD simulation, substructures of which have been used for a comparison of (partly) frozen-density embedding and supermolecule calculations.

TABLE II. Excitation energies (SAOP/TZP/DZ, LDA/DZ for the frozen part; energies in units of eV) of the $n \rightarrow \pi^*$ and $\sigma \rightarrow \pi^*$ valence excitations of acetone in water using the QM/QM embedding scheme for a snapshot from a CPMD simulation, from which subsystems with different numbers n of water molecules have been extracted. Scheme A, supermolecule calculations; scheme B, the two nearest water molecules are included in the embedded system, the other $n-2$ water molecules are frozen; scheme C, all water molecules are frozen. We also show the values for the optimized isolated structure for comparison. For the largest cluster with $n=88$, the frozen density was constructed as a superposition of molecular densities, see text. For the cluster with $n=57$, both methods to construct the density are compared.

	$n \rightarrow \pi^*$			$\sigma \rightarrow \pi^*$		
n	A	B	C	A	B	C
0 (opt.)	4.58	⋯	⋯	7.56	⋯	⋯
0	4.45	⋯	⋯	7.40	⋯	⋯
2	4.57	4.57	4.69	7.09	7.09	7.15
5	4.50	4.54	4.68	7.16	7.15	7.22
10	4.60	4.66	4.74	7.20	7.13	7.22
20	4.71	4.68	4.78	7.19	7.20	7.25
57	⋯	4.72	4.78	⋯	7.14	7.24
57 (mol. frags, diag.)	⋯	4.72	4.79	⋯	7.15	7.24
88 (mol. frags, diag.)	⋯	4.71	4.78	⋯	7.26	7.28

TABLE III. Shifts in the excitation energies compared to the optimized structure of isolated acetone, calculated from the excitation energies obtained with schemes A, B, and C shown in Table II.

	$n\rightarrow\pi^*$			$\sigma\rightarrow\pi^*$		
n	A	B	C	A	B	C
0	−0.13	⋯	⋯	−0.16	⋯	⋯
2	−0.01	−0.01	+0.11	−0.47	−0.47	−0.41
5	−0.08	−0.04	+0.10	−0.40	−0.41	−0.34
10	+0.02	+0.08	+0.16	−0.36	−0.43	−0.34
20	+0.13	+0.10	+0.20	−0.37	−0.36	−0.31
57	⋯	+0.14	+0.20	⋯	−0.42	−0.32
57 (mol. frags)	⋯	+0.14	+0.19	⋯	−0.41	−0.32
88 (mol. frags)	⋯	+0.13	+0.20	⋯	−0.30	−0.28

the *optimized structure* of the isolated ("gas phase") acetone molecule are shown in Table III. These shifts cannot directly be used in a comparison to experiment; a statistical averaging over many snapshots from a (CP)MD simulation is necessary for this purpose, see Sec. V.

A. Charge-transfer excitation problem

As has been explained in Sec. II, the supermolecule calculations are not directly comparable to the embedding calculations presented here. Nevertheless, it has been shown in Ref. 29 that a very good agreement between supermolecule and embedding calculations can be achieved for systems with hydrogen bonds. Comparing embedding (scheme C) and supermolecule calculations (scheme A) in our case, however, reveals differences ranging from 0.02 to 0.18 eV. With one exception, the results obtained with two explicitly nonfrozen water molecules (scheme B) are in better agreement with the supermolecule Kohn–Sham calculations than scheme C, the deviations being between 0.01 and 0.07 eV. Typically, the agreement between calculations with (partly) frozen environment and supermolecule calculations improves if a polarization of the frozen density is allowed by performing freeze-and-thaw cycles.[38] In our case, however, a relaxed density in the calculation (scheme C) for the cluster with two water molecules *increases* the energy of the $n\rightarrow\pi^*$ excitation from 4.69 to 4.73 eV, and for the cluster with 57 water molecules from 4.78 to 4.84 eV, so that the differences are getting even larger. We will explain in the following why the supermolecule and embedding calculations do show these disagreements, and why the embedding calculations are more appropriate for the system under study.

We cannot consider the supermolecule calculations of scheme A as the benchmark values by which to judge the embedding scheme; serious practical problems arise in the identification of certain excitations if (part of) the water molecules are included in the nonfrozen region (i.e., particularly in scheme A, but to a certain extent also in scheme B): In many cases, there is a $n_{\mathrm{O(water)}}\rightarrow\pi^*$ charge-transfer excitation from one of the explicitly included water molecules to the π^* orbital of the acetone, which is apparently too low in excitation energy (cf. the discussion in Refs. 20 and 21). This is one instance of the well recognized problem of too low charge-transfer excitations in TDDFT calculations with the standard LDA or GGA exchange-correlation kernels.[22–24] If only two nonfrozen water molecules are used (scheme B), it is easily possible to identify the $n\rightarrow\pi^*$ excitation, but in some cases, there is already a pronounced mixing between different molecular orbital (MO) contributions to these transitions, and also some mixing between lone pairs of the carbonyl oxygen with those of nearby water oxygens. For the supermolecule calculation in scheme A this mixing is even worse. In order to analyze this problem, we have calculated the overlap of (modified) transition densities, based on TDDFT solution vectors for the system in which water molecules are included in the nonfrozen part (scheme A or B), with those of the corresponding reference calculations of either the isolated molecule or the fully embedded molecule (scheme C). This overlap can be used as a criterion to determine the excitation of interest within the dense manifold of excitations obtained, in particular, in the supermolecule calculation. The overlap $S^{\mathrm{td}}_{AA'}$ for excited states A and A' of the two different systems is calculated in terms of those atomic orbital (AO) contributions which are present in both systems,

$$S^{\mathrm{td}}_{AA'}=\sum_{\mu\nu\lambda'\rho'} X^{\mathrm{AO},A}_{\mu\nu}X^{\mathrm{AO},A'}_{\lambda'\rho'}S^{\mathrm{AO}}_{\mu\lambda'}S^{\mathrm{AO}}_{\nu\rho'}, \tag{9}$$

where μ,ν,λ',ρ' label atomic basis functions for the (partly) explicitly solvated system (unprimed) and the reference system (primed). The sum is over only those basis functions present in both calculations. $S^{\mathrm{AO}}_{\mu\lambda'}$ are overlap integrals over AO basis functions for the two different systems,

$$X^{\mathrm{AO},A}_{\mu\nu}=\sum_{i,r} X^{\mathrm{MO},A}_{ir}c_{\mu i}c_{\nu r} \tag{10}$$

are the TDDFT solution vectors transformed to the AO basis, and $c_{\mu i}$ are the MO coefficients for occupied (i) and unoccupied (r) orbitals. This approach is very similar to the identification of states with similar electronic character used to define a diabatizing scheme in Refs. 61 and 62. This leads to a quantitative measure of the similarity between two excitations in cases of mixings between MOs and MO contributions to these excitations.

If we take the substructure with two water molecules as an example, we find—because of the particular orientation between water and acetone—a pronounced mixing between the carbonyl-oxygen lone pairs and the lone pairs of the neighboring water-oxygen atoms, so that there are three MOs with considerable character of the carbonyl oxygen lone pair, which is the highest occupied molecular orbital for isolated acetone. In the embedding calculation (scheme C), however, no such mixing can occur. In the latter case, there is one excitation which has an overlap S^{td} with the $n\rightarrow\pi^*$ excitation for the isolated molecule of 0.998, i.e., the embedding causes no problem with respect to the labeling as $n\rightarrow\pi^*$. However, the $n\rightarrow\pi^*$ excitation in the supermolecule calculation (scheme A) shows a pronounced mixing with $n(\mathrm{OH_2})\rightarrow\pi^*$ excitations, so that the overlap with the isolated molecule excitation is only 0.909, which corresponds to a contribution of only 82.6%. On the other hand, there is an additional excited state at 5.29 eV with a 9.1% contribution

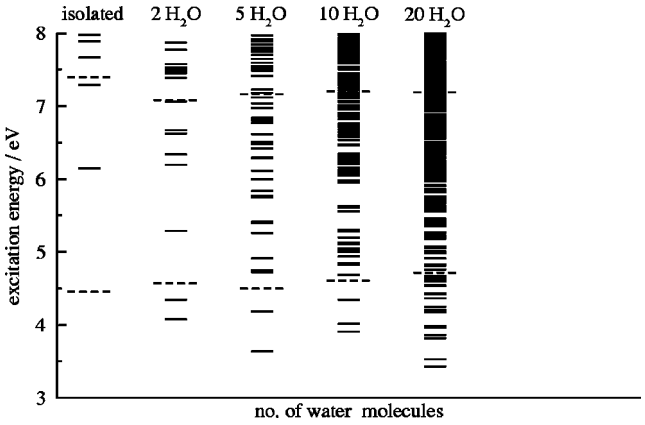

FIG. 2. Excited states (SAOP/TZP/DZ) of acetone-water complexes with different numbers of water molecules. All water molecules have been explicitly taken into account in the TDDFT calculation (scheme A). The longer, dashed lines correspond to the $n\rightarrow\pi^*$ and $\sigma\rightarrow\pi^*$ valence transitions.

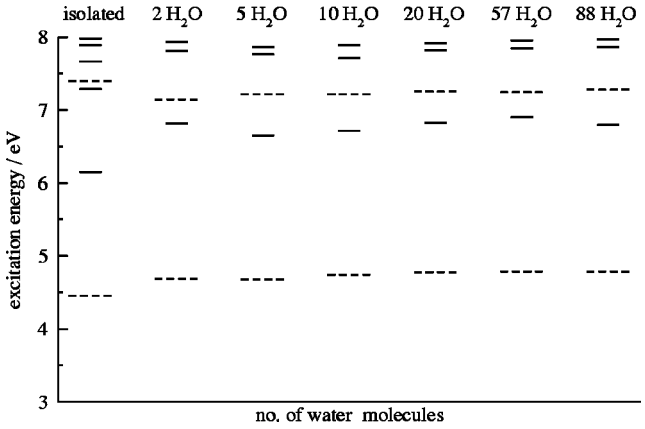

FIG. 4. Excited states (SAOP/TZP/DZ) of acetone-water complexes with different numbers of water molecules. All water molecules are included via the QM/QM embedding scheme in the TDDFT calculation (scheme C). The longer, dashed lines correspond to the $n\rightarrow\pi^*$ and $\sigma\rightarrow\pi^*$ valence transitions.

of the "isolated" $n\rightarrow\pi^*$ excitation (i.e., an overlap of 0.302) and several others with smaller contributions.

This indicates that there arises an ambiguity in the comparison of excitation energies from embedding with supermolecule calculations. The supermolecule TDDFT calculation produces mixings with other MO contributions, which are similar in energy. While this is, in principle, a physical effect that should be included in the calculation, it becomes unphysical in the present example, since the additional excitations are *artificially* too low in energy because of the problem of TDDFT to deal with CT excitations. Solvent to solute transitions should not occur at this low energy and therefore should not interfere with the $n\rightarrow\pi^*$ excitation, as indeed they cannot in the embedding calculation.

This problem becomes more obvious from the graphical representation of the excitation energies for the water clusters shown in Fig. 2 (scheme A), 3 (scheme B), and 4 (scheme C). While scheme C only produces the interesting intrasolute excitations, so that the two valence excitations under study here are always among the lowest four excitations, scheme B already incorporates some excitations from solvent orbitals at low energies, but still the two excitations of interest ($n\rightarrow\pi^*$ and $\sigma\rightarrow\pi^*$) can always be found among the lowest ten excitations. The additional low-lying excitations found here concern mainly excitations from oxygen lone pairs of the nonfrozen water molecules to orbitals of acetone, e.g., $n(OH_2)\rightarrow\pi^*$ excitations. The two additional excitations for the isolated molecule below 7.5 eV (Fig. 4) are of Rydberg type; they are pushed to higher energies by interaction with the solvent molecules. In scheme A, it gets more and more complicated, with increasing number of water molecules, to identify the interesting excitations: For the largest cluster considered in a supermolecule calculation (20 water molecules), the second valence transition cannot be found among the lowest 150 excitations. Moreover, the mixing of the spurious CT excitations with the solute excitation leads to shifts in the excitation energy of the latter, impeding the use of these supermolecule calculations as benchmark for the embedding calculations.

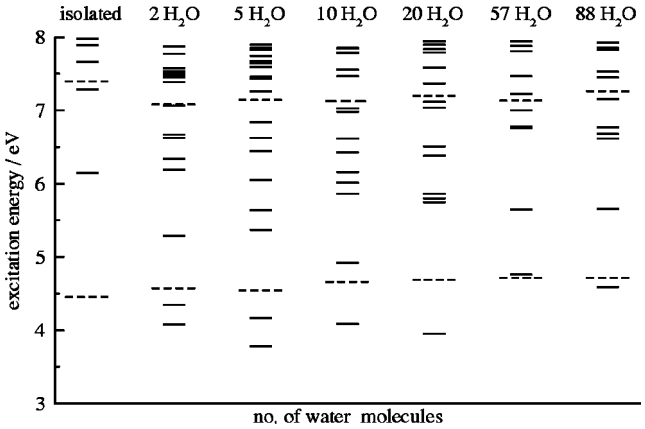

FIG. 3. Excited states (SAOP/TZP/DZ) of acetone-water complexes with different numbers of water molecules. Two water molecules have been explicitly taken into account in the TDDFT calculation (except for the isolated molecule calculation), while all additional water molecules are included via the QM/QM embedding scheme (scheme B). The longer, dashed lines correspond to the $n\rightarrow\pi^*$ and $\sigma\rightarrow\pi^*$ valence transitions.

B. Efficiency of the KSCED excitation energy calculations

A major advantage of the frozen-density embedding lies in the computational efficiency: A supermolecule calculation of the 10 lowest excited states for the cluster with 20 water molecules took 1340 CPU minutes on a SGI origin 3800, while the corresponding embedding calculation (scheme C) took only 311 CPU minutes. In the frozen-density calculations, the time for the construction of the frozen-density in a single-point calculation of the surrounding water molecules is already included in the computational time. Besides the fact that a TDDFT calculation is by far more expensive for the supermolecule calculation anyway, the situation becomes even worse since many more excited states have to be calculated: with only 20 water solvent molecules the $n\rightarrow\pi^*$ excitation is not among the 10 lowest excitations. An additional calculation of 150 excited states showed that it is the 20th excited state. This calculation took 5689 CPU minutes,

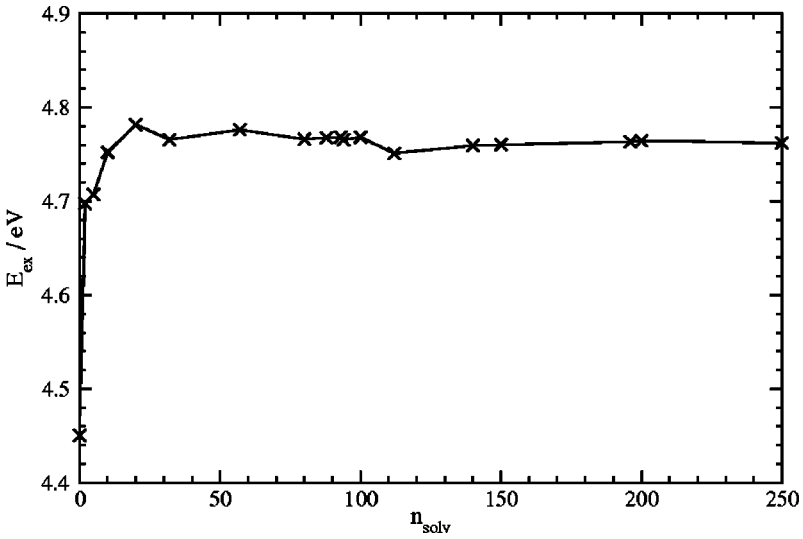

FIG. 5. Excitation energies (SAOP/TZP, LDA/DZ for the frozen part; energies in units of eV) of the $n\rightarrow\pi^*$ valence excitations of acetone in water using the QM/QM embedding scheme for a snapshot from a CPMD simulation, from which subsystems with different numbers n of water molecules have been extracted. The frozen density in all cases was constructed as a sum of densities of water fragments. For clusters with more than 100 water molecules, rigid water molecules where assumed for all but the 20 water molecules nearest to the embedded acetone. For the largest clusters, the number n also includes some acetone molecules from neighboring cells of the CPMD simulation.

about 18 times longer than the embedding calculation. And still the $\sigma\rightarrow\pi^*$ valence excitation is not obtained.

Approximately 4760 CPU minutes were needed to solve the TDDFT equations in this case, compared to about 125 CPU minutes in the embedding calculation. The $\sigma\rightarrow\pi^*$ valence excitation could finally be found in a supermolecule calculation of the lowest 350 states as state no. 216, which took 8517 CPU minutes. Within scheme B, 10 excited states are enough to find both valence transitions, with a computational cost of 349 CPU minutes, of which 135 CPU minutes were needed for the TDDFT part.

C. Convergence with the number of solvent molecules

The data in Table III also allow to draw some conclusions about the appropriate size of the solvation shell necessary to converge the shifts in the excitation energies with respect to the number of solvent molecules, and to determine the importance of different factors on these shifts. We restrict ourselves to a discussion of the calculations of scheme C here, since supermolecule calculations of excitation energies for the largest clusters could not be carried out and the excitation energies are affected by the mixing with CT excitations when nonfrozen water molecules are present; the calculations according to scheme B with two nonfrozen waters show the same general trend as the calculations with full embedding, but in some cases, e.g., the $\sigma\rightarrow\pi^*$ excitation for the $n=20$ cluster, they also show strong mixings.

For both valence excitations in this snapshot there is already a pronounced effect on the excitation energies of just the change of the structure of the acetone: For the isolated acetone molecule with the structure from this snapshot, they decrease by 0.13 eV ($n\rightarrow\pi^*$) or 0.16 eV ($\sigma\rightarrow\pi^*$) compared to the optimized geometry of the isolated molecule. The first two water molecules that are considered in the calculation are bound to the carbonyl oxygen atom by hydrogen bonds. They also have an important effect through their specific interactions: The energy for the $n\rightarrow\pi^*$ excitation increases by 0.24 eV, while for the $\sigma\rightarrow\pi^*$ transition it decreases by 0.25 eV. Further water molecules again cause a blueshift in the excitation energies for both transitions by approximately 0.1 eV. While the excitation energies vary a lot for small numbers of water molecules, the results do not show considerable deviations between the calculations with 20 or 57 water molecules. The latter structure was obtained by cutting a spherical box with a radius of 8.0 Å from the smaller but periodic box of the simulation; this approach was used for all subsequent calculations in the statistical sampling. The additional test with the more approximate density and 88 water molecules shows no change at all for the $n\rightarrow\pi^*$ excitation, while a slight increase of less than 0.04 eV can be observed for the $\sigma\rightarrow\pi^*$ excitation, which is very sensitive to solvent effects anyway. It is instructive to see that there is a shift of about 0.11 eV for scheme B when increasing the number of water molecules from 57 to 88. Again, we attribute this to mixings with excitations involving solvent molecule orbitals: The contribution of the $n\rightarrow\pi^*$ excitation of the isolated system to that of the system with 57 water molecules is only 52.6%, but it increases to 57.1% when the additional 31 water molecules are included. This shows how sensitive these mixings are with respect to slight changes in the orbitals of the explicitly included water molecules.

Even larger numbers of molecules can be treated by this method if, as an additional approximation, the same frozen electron density is used for each water fragment, which appears to be reasonable, in particular, for those molecules which are further away from the solute. A test for the cluster with 88 water molecules described above, in which such a rigid molecule approach for the outer 68 water molecules was used, while the snapshot water structures were employed for the 20 nearest water molecules, resulted in changes lower than 0.003 eV for the valence transitions investigated here. With this approach, and using a simple sum-of-fragments density for the frozen part, we were able to extend the number of solvent molecules to 250. The excitation energies as a function of the number of water molecules are shown in Fig. 5; the rigid molecule approach for the outer water molecules as described above was used for all structures with more than 100 water molecules. These structures were obtained from the same CPMD snapshot, which means that for the largest clusters also acetone molecules in neighboring cells are included in the frozen density. The maximum number of atoms

in this calculation is 802. Again, it can be seen that already with 20 water molecules a plateau is reached, and the changes with increasing numbers of water molecules are very modest, i.e., in the order of 0.01 to 0.02 eV.

This discussion shows that the embedding calculations have great advantages compared to normal Kohn–Sham calculations in the context of solvent effects: (i) The excitation energies are not affected by mixings with spuriously too low CT excitations from solvent to solute (or vice versa), (ii) the calculations are much more efficient, so that larger numbers of water molecules can be described on a quantum chemical footing, and (iii) the interpretation of the results is much easier, since only excitations within the embedded system are obtained.

V. (CP)MD SIMULATIONS

The purpose of the calculations presented in Sec. IV was a comparison of the frozen-density approach to supersystem calculations, demonstrating the problem of the identification of certain valence transitions among spuriously too low charge-transfer excitations in the latter with increasing numbers of water molecules in the calculation. In order to describe the solvatochromic shifts, however, it is necessary to work with a larger ensemble of solvent configurations for a statistical analysis. We thus performed CPMD and classical MD simulations of acetone in water; while snapshots taken from the CPMD simulations are certainly more representative in terms of the structures obtained, the classical MD simulations allow to use much larger periodic box sizes to model the solvation, which is certainly necessary if this method shall be applied to larger solvated molecules.

CPMD simulations have been carried out with the projector augmented wave (PAW) package[58] using the BLYP functional.[48,59] For efficiency reasons, we employed deuterium isotopes for the hydrogen atoms. We used a fictitious mass for the wave functions of 100 a.u., a time step of 6.5 a.u. (0.157 fs), and a small friction on the wave functions of 0.000 05. This has been tested to give good simulations of liquid water. The cutoff for the plane-wave basis was 30 Ry (408 eV) and 90 Ry (1225 eV) for the charge density. The simulation was performed with one acetone molecule and 31 water molecules in a periodic cubic box of 10.168 Å. A Nosé thermostat with an oscillation frequency of 60 000 a.u. was used to keep the temperature at 300 K. After equilibration for 11 ps with more strongly coupled thermostats, the simulation was prolonged for 8 ps. Every 36 fs, a snapshot for the final statistics was taken (220 in total), while general analyses are carried out with snapshots taken in intervals of 182 fs (44 in total).

Additionally, we employed a purely classical model for acetone in water. Parameters from the GAFF[54] force field have been used for acetone, combined with scaled partial charges from a Hirshfeld charge analysis for the isolated acetone molecule (SAOP/TZP). In order to verify this model, we compared the results of our MD simulations with those of the CPMD simulation with the same periodic box of 10.168 Å with one acetone molecule and 31 water molecules. The O(carbonyl)–H(water) and O(carbonyl)–O(water) radial distribution functions from both simulations are shown in Fig. 6. The partial charges for the acetone molecule have been uniformly scaled in order to obtain optimum agreement of these radial distribution functions (scaling factor: 1.61). With this classical model, we performed simulations at 300 K of a periodic cubic box of 21.225 Å with 315 TIP3P water molecules and one acetone molecule. After a structure optimization and an equilibration phase of 200 ps (with 2.0 fs time steps), we carried out a simulation of 50 ps. Every 2 ps, a snapshot of the box was taken.

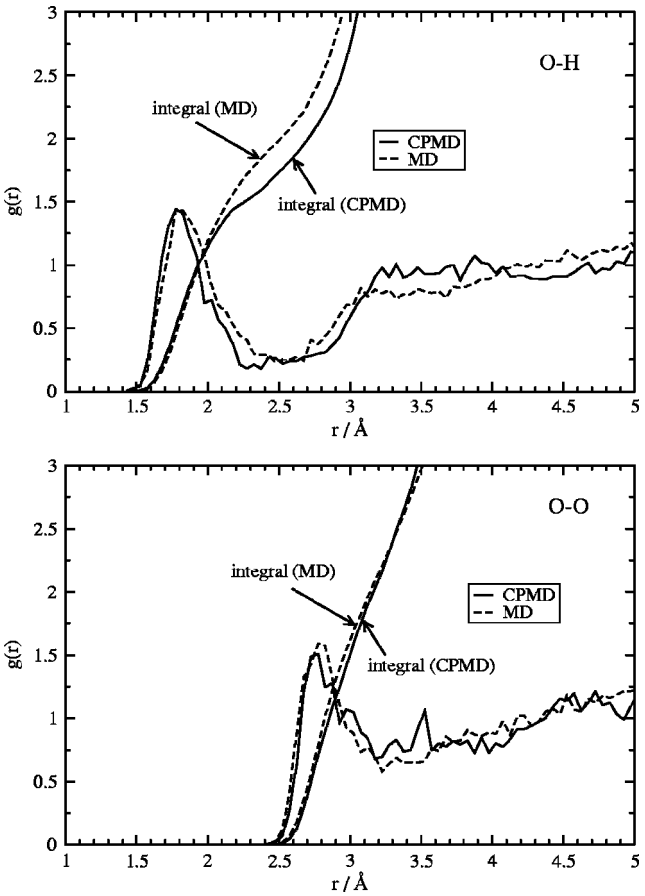

FIG. 6. Radial distribution function $g(r)$ and integral from a MD simulation (GAFF/TIP3P) in comparison to a CPMD simulation of the same system. Top, O(carbonyl)–H(water) distribution function; bottom, O(carbonyl)–O(water) distribution function.

In Fig. 6 the radial distribution functions around the carbonyl oxygen are compared. After scaling the partial charges for the acetone molecule especially the first peaks are very well reproduced by the classical MD simulations. This means that the distances and the number of hydrogen bonds to the acetone oxygen are correctly modeled. Minor difference is that the first peak in the O(carbonyl)–H(water) radial distribution function has a bit of a tail and at longer distances the peaks are slightly less pronounced. The orientation of the hydrogen bonds around the carbonyl group is much less well reproduced by the classical simulations. In the CPMD simulations the hydrogen bonds are in the plane of the acetone molecule (dihedral angles with the methyl groups: $0° \pm 30°$) with angles around $150° (\pm 10°)$ with the CO bond. In the classical simulations this chemical preference for certain

angles was not present at all. Naturally, this could affect the solvent shift calculated from these structures.

For all following calculations, only the acetone molecule was explicitly considered, while the density of all surrounding water molecules was frozen (scheme C). Considering all water molecules within a sphere with radius $r=8$ Å as explained in Sec. IV, we obtained systems with between 50 and 61 water molecules from the (CP)MD snapshots, for which we calculated vertical excitation energies using the SAOP potential. A TZP basis set was used for acetone, while a DZ basis set was used for the water molecules in the calculation of the solvent density, which was obtained within the LDA. Note that on average about 175 atoms are included in the QM/QM TDDFT calculations.

The optical spectrum $f(\omega)$ of acetone was calculated through,[25]

$$f(\omega) = \frac{1}{N_c} \sum_{i=1}^{n_{\max}} \sum_{j=1}^{N_c} f_i^j \delta(\omega_i^j - \omega), \tag{11}$$

where N_c is the number of solvent configuration snapshots taken into account, $n_{\max}$ is the number of excitations considered for each configuration, ω_i^j is the vertical excitation frequency for a particular state and configuration, and f_i^j is the corresponding oscillator strength.

There are different ways to extract the solvent shift from the calculation, which are more or less well suited in the present case. The simplest way is just to calculate the difference between average transition energies E_{avg} for the gas phase and for the solvated molecule (so averaging over the snapshot configurations without any weighting). Another possibility is to use approximate mean transition frequencies (or the corresponding mean transition energies E_{mte}) for a particular excitation i, in which the vertical transition energies of the snapshots are weighed with the oscillator strengths (cf. Ref. 21),

$$\langle \omega_i \rangle \approx \frac{\sum_{j=1}^{N_c} f_i^j \omega_i^j}{\sum_{j=1}^{N_c} f_i^j}. \tag{12}$$

These values are probably more appropriate because they emphasize those parts of the spectrum with higher intensities. Experimentally, the solvent shifts are extracted by taking the positions of maximum intensity in the spectrum, $E_{\max}$, which can of course also be extracted from the spectrum simulation. In contrast to E_{mte}, these values do not directly depend on the shape of an absorption band; symmetric and nonsymmetric band shapes may still have the same maximum position. However, taking the position of maximum intensity in the simulated spectrum is problematic as long as the number of configurations is small, since single solvent configurations accidentally leading to high oscillator strengths can easily shift the maximum position considerably. It has recently been observed[30] that several thousand's of configurations might be necessary to model a complete absorption band shape correctly. Therefore, the $E_{\max}$ values are not very reliable if it is not feasible to sample such a large number of configurations, and it seems more appropriate in such a case to extract the maximum position of a Gaussian fit to the simulated spectrum, $E_{\max,\text{fit}}$. Hence, we consider this last way the most appropriate choice in our study for a comparison of our simulated data with the results extracted from experiment.

We carried out general analyses with a small number of snapshots for different ways to describe the solvent density. For each of these approaches, we present the calculated energies and solvent shifts extracted with all four different methods described above in Table IV. A gas phase CPMD simulation for isolated acetone at 300 K resulted in a value for $E_{\max,\text{fit}}$ of 4.4746 eV obtained from 300 snapshots of a 12 ps simulation time (compared to a vertical excitation energy of 4.58 eV for the optimized structure), and all other ways to estimate the band maximum agree within 0.019 eV. This estimate for the peak maximum is in very good agreement with the experimental value of 4.48 (Refs. 32 and 34)–4.49 eV.[33] Note that in the gas phase it is not possible to use a conventional thermostat, so the temperature in this CPMD simulation was kept constant by applying a very small negative friction to the atoms.

As a starting point for the solvation models, we used 44 solvent structures from the CPMD simulation as described above. We first applied the default scheme to obtain the frozen density, i.e., all water molecules were frozen, and the density was obtain in an SCF calculation on the full set of water molecules with a convergence criterion $s_{\text{conv}}=0.1$ (see Sec. III). This results in solvent shifts between 0.2150 (based on E_{avg}) and 0.2493 eV (based on E_{mte}), so that all shifts agree within 0.034 eV. In order to show that the frozen density can also be constructed by a superposition of molecular densities as described in Sec. III, we calculated the solvent shifts for the same set of structures using this approach for the density of the environment. The changes are very slight if one explicit Fock-matrix diagonalization is carried out ("diag." in Table IV), with solvent shifts between 0.2078 and 0.2419 eV. The deviations are slightly larger, but still acceptably small, with solvent shifts between 0.1995 and 0.2324, if a simple superposition of water densities is performed ("sumf." in Table IV). As a next test for the same set of structures, we first calculated the density of the acetone molecule for each structure, which was then frozen to calculate a relaxed density of the surrounding solvent. This relaxed solvent density was then employed in an embedding calculation for the acetone. The relaxed environment leads to solvent shifts that are a bit higher, typically in a quite systematic way by about 0.06 eV. This is a case where the maximum in the simulated spectrum is "accidentally" too high, leading to an anomalously large shift of 0.3920. The other shifts are around 0.30 eV, and agree to within 0.038 eV. These results are somewhat out of line, although we expect a significant effect (i.e., a decrease by about 0.04 eV) of using an extended set of solvent configurations (see below). Nevertheless, further work is envisaged, involving additional freeze-and-thaw cycles and incorporation of basis sets on the frozen-density subsystems to determine if this is a converged result in the sequence of alternating system I–system II relaxations.

In an additional test, which is not shown in Table IV, we included the two nearest water molecules explicitly in the

TABLE IV. Estimations of the $n \rightarrow \pi^*$ solvent shift of acetone in water (compared to the vapor spectrum) using different ways to estimate the peak maxima. All energies given in units of eV. Also given is the number of configurations $n_{\text{conf.}}$ used to extract the values. The labels have the following meanings: mte, mean transition energy; avg, average transition energy; max, peak maximum in simulated spectrum; max, fit, peak maximum in Gaussian fit to simulated spectrum (recommended).

Structure	ρ_{H_2O}		$n_{\text{conf.}}$	E_{mte}	E_{avg}	E_{max}	$E_{\text{max,fit}}$
Vapor							
CPMD	···		300	4.4601	4.4740	4.4560	4.4746
Expt.[a]							4.48 to 4.49
Solution							
CPMD	Frozen ($s_{\text{conv}}=0.1$)		44	4.7094	4.6890	4.6940	4.7221
		Shift		0.2493	0.2150	0.2380	0.2475
CPMD	Frozen (mol. frags., sumf.)		44	4.6925	4.6735	4.6700	4.7131
		Shift		0.2324	0.1995	0.2140	0.2385
CPMD	Frozen (mol. frags., diag.)		44	4.7020	4.6818	4.6940	4.7149
		Shift		0.2419	0.2078	0.2380	0.2403
CPMD	Frozen ($s_{\text{conv}}=0.1$), relaxed		44	4.7686	4.7449	4.8480	4.7832
		Shift		0.3085	0.2709	0.3920	0.3086
MD	Frozen ($s_{\text{conv}}=0.1$)		25	4.6793	4.6793	4.7240	4.6963
		Shift		0.2192	0.2053	0.2680	0.2217
CPMD	Frozen ($s_{\text{conv}}=0.1$)		220	4.7184	4.7000	4.6740	4.6791
		Shift		0.2583	0.2260	0.2180	0.2045
CPMD	Frozen ($s_{\text{conv}}=0.1$)		219	4.6883	4.6978	4.6740	4.6791
		Shift		0.2282	0.2238	0.2180	0.2045
Expt.[b]							4.68–4.69
		Shift					0.19–0.21

[a]References 32–34.
[b]References 32–36.

embedded system, again using the same 44 snapshots. Compared to the first simulation, this typically reduces the solvent shift by about 0.04 eV, leading to values between 0.1739 eV (based on E_{avg}), and 0.2255 eV (based on $E_{\text{max,fit}}$). The value based on E_{max} again is an exception; with 0.3240 eV it is much higher than all other shifts based on this model for the solvation. This test suffers, however, from the mixings mentioned in earlier sections of this work: For practically every snapshot, more or less strong mixings with $n(\text{OH}_2) \rightarrow \pi^*$ excitations could be observed. In many cases, the identification of the $n \rightarrow \pi^*$ excitation was only possible based on the quantitative measure given in Eq. (9); some snapshots gave rise to two excitations with competing $n \rightarrow \pi^*$ contributions, which shows that the resulting statistics is not very reliable.

A last comparison was made with 25 structures obtained from a classical MD simulation. The solvent shifts are between 0.2053 (based on E_{avg}) and 0.2217 (based on $E_{\text{max,fit}}$); also here, the E_{max} value is somewhat of an outlier with 0.2680, which demonstrates again that it is a less reliable measure of the shift. No large discrepancies are therefore observed in comparison to the CPMD simulation. However, the structures from two of the snapshots lead to severe convergence problems in the SCF calculation for the solvent because of near-degenerate orbitals at the HOMO-LUMO gap, and subsequently to unreliable TDDFT results. This seems to be purely a problem of the SCF procedure, caused by an unrealistic structure of the solvent: If only the superposition of the solvent fragment densities is used (scheme "diag."), the resulting excitation energies and oscillator strengths are within the normal range. Therefore, we have to conclude that the classical MD occasionally leads to structures which can induce SCF problems, so that the resulting densities are not in all cases reliable approximations. In other studies, such problems with classical MD simulations are circumvented by performing a small CPMD relaxation study after a QM/MM CPMD study.[5]

The final spectrum was simulated using 220 solvent configurations from the CPMD simulation, in which the total water environment density was frozen and unrelaxed; the density was obtained from an SCF calculation with a convergence threshold of $s_{\text{conv}}=0.1$. The band of the $n \rightarrow \pi^*$ transition is shown in Fig. 7, together with the simulation of gas-phase acetone. The solvent shift extracted from the Gaussian fit to the band is 0.2045 eV. This is 0.0430 eV lower than for the smaller set of 44 snapshots, and in excellent agreement with the experimental values of 0.19–0.21 eV and with former theoretical studies mentioned in the introduction. Also the maximum position of the spectrum and the average transition energies confirm this value, differing by less then 0.02 eV in the calculated shift. Only the mean transition energy results in a somewhat higher shift of 0.2583 eV, which can still be considered a good estimate for the experimental value. This somewhat too high value can exclusively be attributed to one outlier in our data set: we obtained one configuration with an excitation energy of 5.16, which has a very high oscillator strength and is therefore responsible for the peak at the high-frequency border of this band. If we neglect this one outlier, the shift computed by mean transition energies decreases to 0.2282 eV, in much better agreement with

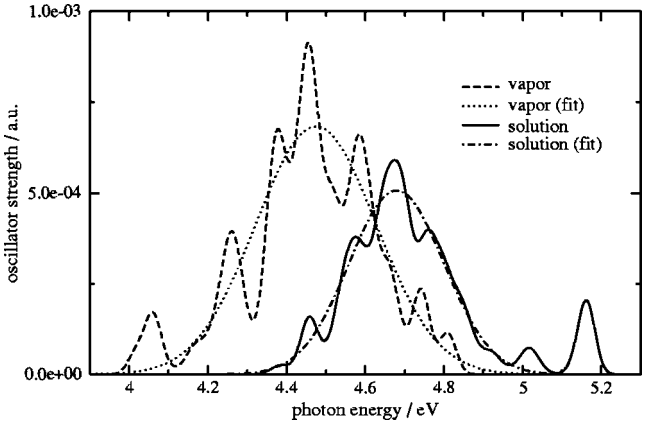

FIG. 7. Simulated (SAOP/TZP/DZ) absorption spectrum of acetone in water. Excitation energies have been calculated for snapshots from a CPMD simulation for acetone vapor (dashed line) or acetone in water (solid line), respectively; all water molecules within a radius of 8 Å from the acetone molecule have been considered in the latter calculations, and have been treated in a frozen-density fashion. In total, 300 (vapor) or 220 (solution) configurations have been sampled. A Gaussian smearing of 0.05 eV has been applied to the peaks in the spectrum. Additionally, the curves of a Gaussian fit to the simulated absorption bands are shown for the gas-phase (dotted line) or solvated (dashed-dotted line) molecule, respectively.

the other values. The maximum position of the spectrum does not change at all, the average transition energy changes only very little, and also the changes in the maximum of the fit are negligible. This latter point is due to the fact that we use a symmetric Gaussian function to fit our data, so that a better description of the outlier in the fit would dramatically decrease the quality of the fit in the low-frequency regime of this band. It therefore does not affect this fit very much. We would like to note that not only the solvent shift but also the estimated band maximum of 4.6791 eV obtained in this way is in very good agreement with the experimental values of 4.68 (Ref. 33)–4.69 eV.[32,34–36]

Our tests with relaxed frozen densities showed that relaxation typically leads to an increase of the excitation energies of about 0.06 eV, which was quite systematic for the structures investigated. Therefore, we expect that relaxation would also increase the solvent shift for the large set of 220 snapshots to about 0.26 eV. The solvent shifts obtained with different ways to construct the frozen density can therefore be estimated to lie in the range of 0.20–0.26 eV.

There is another interesting point about our simulation: In the experimental spectra in Refs. 32 and 33, it can be seen that the structureless vapor absorption band of the $n \rightarrow \pi^*$ transition is a bit broader than the spectrum in aqueous solution. Although it is difficult to extract exact values, the half-widths can be estimated to roughly 7000 cm^{-1} (0.87 eV vapor) and 6000 cm^{-1} (0.74 eV, solution). The Gaussian fits to our final simulated absorption bands qualitatively agree with this observation; while the fitted gas-phase spectrum has a half-width of 0.51 eV, the solution spectrum has a reduced half-width of only 0.40 eV. But in terms of absolute values, these half-widths are still too low, and a much larger statistics might be necessary to draw definite conclusions about the full shape of the absorption band. The calculated absorption intensities are less reliable; the basis set requirements for converging them are usually much larger than for the excitation energies. Moreover, the confinement of the response to the embedded system may make the electron density of the embedded system less polarizable. We therefore refrain from an interpretation of changes in the intensities in terms of physical effects in this work; further investigations on this point are in progress.

VI. CONCLUSION

The calculations in this work clearly show the potential of the frozen density KSCED embedding in the calculation of excitation energies for solvated systems. In the first place, this embedding approach implies a restriction to the interesting intrasubsystem excitations, where the embedded system can be defined by the user. Other excitations, which might be real or spuriously induced by failures of the (TD)DFT approach are not obtained. This has a number of important consequences for the calculation of solvated systems: (i) Embedding calculations are inherently less demanding, especially for TDDFT calculations with many solvent molecules. (ii) In contrast to supermolecule calculations on systems like the one studied here, the number of low-lying excitations is constant, which makes embedding calculations even more advantageous when large solvent shells are included. (iii) The interpretation of the results is much easier, and the excitations are not affected by mixings with spuriously low CT excitations.

It could be shown that the excitation energies obtained from the embedding scheme are not very sensitive to approximations made in the construction of the frozen density. This is an important observation, since otherwise the determination of the frozen density would be the time-limiting step for very large systems. The possibility to use superpositions of molecular densities is especially important, which replaces a calculation on a system with N molecules by N calculations on systems of just one individual molecule. This can be further simplified by assuming a rigid structure of the solvent molecules, at least in the outer solvation shells, so that just one calculation on one molecule is necessary for the preparation of their frozen density. Such simplifications allow to perform statistical analyses of excitation energies for many snapshots of the solvent, which was demonstrated here for systems with on average 175 atoms. Some test calculations in this work demonstrated that calculations with several hundreds of solvent molecules, or more than 800 atoms are feasible with the approximate frozen-density construction schemes.

In combination with snapshots from CPMD simulations, this enabled us to model the solvatochromic shift of the $n \rightarrow \pi^*$ excitation of acetone in water. An estimated shift of 0.20 eV was obtained from a statistical analysis of excitation energies for 220 snapshots of acetone in water. This is in perfect agreement with the experimental shifts of 0.19–0.21 eV.[32–36] However, there is a caveat: In a preliminary study of density relaxation of the solvent system, we found an increase of the calculated solvent shift by about 0.06 eV, so that an estimated solvent shift from the relaxed frozen-density calculations would be in the order of 0.26 eV,

which is somewhat too high, even if uncertainties in the experimental band maxima in the order of 0.01 eV are taken into account. Possible error sources here could be the nonadditive kinetic energy functional, the basis set limited to the embedded system, and the limited statistics of only 44 snapshots for the polarized frozen densities. In view of the variety of factors influencing the simulation of solvatochromic shifts on a purely quantum chemical footing, these results are still very encouraging for further tests and improvements of this approach. The frozen-density embedding will thus allow the calculation of solvatochromic shifts for much larger molecules, provided that reliable structures for snapshots of the solute and its surrounding solvation shells can be obtained.

ACKNOWLEDGMENTS

J.N. gratefully acknowledges funding by a Forschungsstipendium of the Deutsche Forschungsgemeinschaft (DFG). M.J.L. acknowledges a grant from the Dutch National Research School Combination "Catalysis by Design" (NRSC-C). This work was supported by the Swiss National Science Foundation (SNSF).

[1] C. Reichardt, *Solvents and Solvent Effects in Organic Chemistry* (Verlag Chemie, Weinheim, 1988).
[2] C. Reichardt, Chem. Rev. (Washington, D.C.) **94**, 2319 (1994).
[3] E. S. Dodsworth, M. Hasegawa, M. Brigde, and W. Linert, in *Comprehensive Coordination Chemistry*, edited by A. B. P. Lever (Elsevier, Amsterdam, 2004), Vol. 2, 351–365.
[4] A. H. de Vries and P. T. van Duijnen, Int. J. Quantum Chem. **57**, 1067 (1996).
[5] U. F. Röhrig, I. Frank, J. Hutter, A. Laio, J. VandeVondele, and U. Rothlisberger, ChemPhysChem **4**, 1177 (2003).
[6] M. Odelius, B. Kirchner, and J. Hutter, J. Phys. Chem. A **108**, 2044 (2004).
[7] J. Zeng and D. Xie, J. Comput. Chem. **25**, 813 (2004).
[8] D. Xie and J. Zeng, J. Comput. Chem. **25**, 1487 (2004).
[9] W.-G. Han, T. Liu, F. Himo, A. Toutchkine, D. Bashford, K. M. Hahn, and L. Noodleman, ChemPhysChem **4**, 1084 (2003).
[10] T. Liu, W.-G. Han, F. Himo, G. M. Ullmann, D. Bashford, A. Toutchkine, K. M. Hahn, and L. Noodleman, J. Phys. Chem. A **108**, 3545 (2004).
[11] M. Sulpizi, P. Carloni, J. Hutter, and U. Rothlisberger, Phys. Chem. Chem. Phys. **5**, 4798 (2003).
[12] J. Zeng, J. S. Craw, N. S. Hush, and J. R. Reimers, J. Phys. Chem. **98**, 11075 (1994).
[13] S. Hush and J. R. Reimers, Coord. Chem. Rev. **177**, 37 (1998).
[14] N. S. Hush and J. R. Reimers, Chem. Rev. (Washington, D.C.) **100**, 775 (2000).
[15] S. Fantacci, F. D. Angelis, and A. Selloni, J. Am. Chem. Soc. **125**, 4381 (2003).
[16] J. Tomasi and M. Persico, Chem. Rev. (Washington, D.C.) **94**, 2027 (1994).
[17] C. J. Cramer, *Essentials of Computational Chemistry* (Wiley, New York, 2002).
[18] A. Klamt, V. Jonas, T. Bürger, and J. C. W. Lohrenz, J. Phys. Chem. A **102**, 5074 (1998).
[19] R. Car and M. Parrinello, Phys. Rev. Lett. **55**, 2471 (1985).
[20] L. Bernasconi, M. Sprik, and J. Hutter, J. Chem. Phys. **119**, 12417 (2003).
[21] L. Bernasconi, M. Sprik, and J. Hutter, Chem. Phys. Lett. **394**, 141 (2004).
[22] A. Dreuw, J. L. Weisman, and M. Head-Gordon, J. Chem. Phys. **119**, 2943 (2003).
[23] Y. Tawada, T. Tsuneda, S. Yanagisawa, T. Yanai, and K. Hirao, J. Chem. Phys. **120**, 8425 (2004).
[24] O. Gritsenko and E. J. Baerends, J. Chem. Phys. **121**, 655 (2004).
[25] L. Bernasconi, J. Blumberger, M. Sprik, and R. Vuilleumier, J. Chem. Phys. **121**, 11885 (2004).
[26] A. Osted, J. Kongsted, K. V. Mikkelsen, and O. Christiansen, J. Phys. Chem. A **108**, 8646 (2004).
[27] T. A. Wesolowski and A. Warshel, J. Phys. Chem. **97**, 8050 (1993).
[28] M. E. Casida and T. A. Wesolowski, Int. J. Quantum Chem. **96**, 577 (2004).
[29] T. A. Wesolowski, J. Am. Chem. Soc. **126**, 11444 (2004).
[30] N. A. Besley, M. T. Oakley, A. J. Cowan, and J. D. Hirst, J. Am. Chem. Soc. **126**, 13502 (2004).
[31] N. A. Besley, Chem. Phys. Lett. **390**, 124 (2004).
[32] C. W. Porter and C. Iddings, J. Am. Chem. Soc. **48**, 40 (1926).
[33] N. S. Bayliss and E. G. McRae, J. Phys. Chem. **58**, 1006 (1954).
[34] N. S. Bayliss and G. Wills-Johnson, Spectrochim. Acta, Part A **24**, 551 (1968).
[35] W. P. Hayes and C. J. Timmons, Spectrochim. Acta **21**, 529 (1965).
[36] A. Balasubramanian and C. N. R. Rao, Spectrochim. Acta **18**, 1337 (1962).
[37] P. Cortona, Phys. Rev. B **44**, 8454 (1991).
[38] T. A. Wesolowski and J. Weber, Int. J. Quantum Chem. **61**, 303 (1997).
[39] T. A. Wesolowski, J. Chem. Phys. **106**, 8516 (1997).
[40] T. A. Wesolowski, Y. Ellinger, and J. Weber, J. Chem. Phys. **108**, 6078 (1998).
[41] H. Lee, C. Lee, and R. G. Parr, Phys. Rev. A **44**, 768 (1991).
[42] T. A. Wesolowski, H. Chermette, and J. Weber, J. Chem. Phys. **105**, 9182 (1996).
[43] J. P. Perdew, in *Electronic Structure of Solids*, edited by P. Ziesche and H. Eschrig (Akademie Verlag, Berlin, 1991), p. 11.
[44] A. Lembarki and H. Chermette, Phys. Rev. A **50**, 5328 (1994).
[45] M. Dulak and T. A. Wesolowski, Int. J. Quantum Chem. **101**, 543 (2005).
[46] Amsterdam density functional program, Theoretical Chemistry, Vrije Universiteit, Amsterdam, URL: http://www. scm.com
[47] G. te Velde, F. M. Bickelhaupt, E. J. Baerends, S. J. A. van Gisbergen, C. F. Guerra, J. G. Snijders, and T. Ziegler, J. Comput. Chem. **22**, 931 (2001).
[48] A. D. Becke, Phys. Rev. A **38**, 3098 (1988).
[49] J. P. Perdew, Phys. Rev. B **33**, 8822 (1986).
[50] P. R. T. Schipper, O. V. Gritsenko, S. J. A. van Gisbergen, and E. J. Baerends, J. Chem. Phys. **112**, 1344 (2000).
[51] O. V. Gritsenko, P. R. T. Schipper, and E. J. Baerends, Chem. Phys. Lett. **302**, 199 (1999).
[52] O. V. Gritsenko, P. R. T. Schipper, and E. J. Baerends, Int. J. Quantum Chem. **76**, 407 (2000).
[53] M. Grüning, O. V. Gritsenko, S. J. A. van Gisbergen, and E. J. Baerends, J. Chem. Phys. **116**, 9591 (2002).
[54] J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, and D. A. Case, J. Comput. Chem. **25**, 1157 (2004).
[55] W. L. Jorgensen, J. Chandrasekhar, and J. D. Madura, J. Chem. Phys. **79**, 926 (1983).
[56] J. W. Ponder, TINKER—Software Tools for Molecular Design, Version 4.2, http://dasher.wustl.edu/tinker, 2004.
[57] J. W. Ponder and F. M. Richards, J. Comput. Chem. **8**, 1016 (1987).
[58] P. E. Blöchl, Phys. Rev. B **50**, 17953 (1994).
[59] C. Lee, W. Yang, and R. G. Parr, Phys. Rev. B **37**, 785 (1988).
[60] C. F. Guerra, J. G. Snijders, G. te Velde, and E. J. Baerends, Theor. Chem. Acc. **99**, 391 (1998).
[61] J. Neugebauer, E. J. Baerends, and M. Nooijen, J. Chem. Phys. **121**, 6155 (2004).
[62] J. Neugebauer, E. J. Baerends, and M. Nooijen, J. Phys. Chem. A (in press).