Molecular quantum mechanical gradients within the polarizable embedding approach—Application to the internal vibrational Stark shift of acetophenone

Nanna Holmgaard List,[1,a)] Maarten T. P. Beerepoot,[2] Jógvan Magnus Haugaard Olsen,[1,3] Bin Gao,[2] Kenneth Ruud,[2] Hans Jørgen Aagaard Jensen,[1] and Jacob Kongsted[1]
[1]*Department of Physics, Chemistry and Pharmacy, University of Southern Denmark, Campusvej 55, Odense M, Odense DK-5230 Denmark*
[2]*Centre for Theoretical and Computational Chemistry, Department of Chemistry, University of Tromsø–The Arctic University of Norway, N-9037 Tromsø, Norway*
[3]*Laboratory of Computational Chemistry and Biochemistry, Ecole Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland*

(Received 6 November 2014; accepted 2 January 2015; published online 21 January 2015)

We present an implementation of analytical quantum mechanical molecular gradients within the polarizable embedding (PE) model to allow for efficient geometry optimizations and vibrational analysis of molecules embedded in large, geometrically frozen environments. We consider a variational ansatz for the quantum region, covering (multiconfigurational) self-consistent-field and Kohn–Sham density functional theory. As the first application of the implementation, we consider the internal vibrational Stark effect of the C═O group of acetophenone in different solvents and derive its vibrational linear Stark tuning rate using harmonic frequencies calculated from analytical gradients and computed local electric fields. Comparisons to PE calculations employing an enlarged quantum region as well as to a non-polarizable embedding scheme show that the inclusion of mutual polarization between acetophenone and water is essential in order to capture the structural modifications and the associated frequency shifts observed in water. For more apolar solvents, a proper description of dispersion and exchange–repulsion becomes increasingly important, and the quality of the optimized structures relies to a larger extent on the quality of the Lennard-Jones parameters. © *2015 AIP Publishing LLC.* [http://dx.doi.org/10.1063/1.4905909]

I. INTRODUCTION

The organization of charged, polar, and polarizable amino acids in a structured protein matrix generates large electric fields that contribute to virtually all aspects of protein function. Enzymatic catalytic power and ligand binding[1–5] as well as optical properties of embedded chromophores,[6–8] to mention a few, may all be substantially tuned or even driven by the internal electric fields in proteins. While there is no direct way of measuring electric fields in proteins, spectroscopic observables that mirror local electric fields can, once calibrated, act as quantitative reporters of the electrostatic changes occurring upon, e.g., mutations or structural rearrangements of the amino acids.

Vibrational probes have proven to be particularly well suited to act as local and directional metrics of protein electric fields via the vibrational Stark effect,[9–17] which describes the susceptibility of a vibrational transition to an electric field. Often one uses "internal" or "external" to distinguish the source of the Stark effect, being either a surrounding environment or an externally applied field. Protein fields are generally much larger (1–100 MV cm^{-1}) than the uniform fields that can be applied in Stark experiments (~1 MV cm^{-1}) and vary significantly in both magnitude and direction between sites in the protein.[8,10,18,19] Obviously, the spatial resolution that can be obtained is limited by the size of the probe. This renders localized vibrational transitions, such as functional group stretching modes, superior to electronic transitions, which typically extend over the entire molecular (chromophore) skeleton, thereby providing only an average picture of the field within the molecular volume.[19–22] Vibrational probes are smaller than electronic chromophores and thus potentially less perturbative, and the associated transitions also exhibit smaller difference polarizabilities compared to their electronic counterparts. Consequently, when electrostatic effects dominate, many vibrations behave more or less according to the simplified linear Stark effect theory. Within this model, the shift in vibrational frequency ($\Delta\nu$), reflecting a variation in the local electric field ($\Delta\mathbf{F}_{\text{local}}$) experienced by the probe, is assumed to be dictated by the field projection onto the difference dipole vector between the excited and ground vibrational states of the probe $\Delta\boldsymbol{\mu}_{\text{probe}}$

$$hc\Delta\nu = -\Delta\boldsymbol{\mu}_{\text{probe}} \cdot \Delta\mathbf{F}_{\text{local}}, \tag{1}$$

where $\Delta\boldsymbol{\mu}_{\text{probe}}$ is also known as the linear vibrational Stark tuning rate, h is Planck's constant, and c is the speed of light.

Vibrational Stark effect spectroscopy applied to an immobilized, isotropic sample provides the direction and magnitude of the vibrational Stark tuning rate for the vibrational mode of the probe. For approximate one-dimensional modes, e.g., a decoupled stretching between two atoms, $\boldsymbol{\mu}_{\text{probe}}$ will be aligned

a)Electronic mail: nhl@sdu.dk

with the bond axis.[23] With the Stark tuning rate of the probe known, we can apply a reverse strategy by virtue of Eq. (1) to relate a vibrational frequency shift to the concomitant change in the electric field sensed by the probe due to a modification of the surrounding environment. However, the determination of total fields at the probe location requires a frequency–field calibration of the probe on an absolute scale. Whenever a high-resolution structure of the protein is available, computations constitute a complementary tool that allows a direct evaluation of the internal electric fields.[24] Correlating the computed fields with the vibrational frequency of the probe (experimental or calculated), one obtains a calibration curve that provides a quantitative mapping of the electrostatic landscape in other complex environments based on the vibrational frequency of the probe.

The polarizable embedding (PE) model[25,26] is a discrete quantum–classical approach tailored for efficient yet accurate inclusion of the electrostatic and induction effects of a surrounding environment. It has been successfully applied to model the environmental effects on both electric and magnetic molecular properties.[8,27,28] The aim of the present study is to broaden the scope of the PE model to vibrational properties of molecules embedded in large and complex environments. To this end, we have implemented molecular gradients with respect to nuclear displacements in the quantum region within the PE framework. This allows for partial geometry optimizations and normal mode analyses based on numerically calculated Hessians of an embedded molecule described at the Hartree–Fock (HF), Kohn–Sham density functional theory (DFT), or multiconfigurational self-consistent field (MCSCF) level of theory. We use a sequential approach to compute the vibrational properties of an embedded molecule that consists of PE property calculations on statistically independent configurations derived from molecular dynamics (MD) simulations. In comparison to a geometry optimization and Hessian analysis of the full molecular system, this procedure preserves the thermal fluctuations in the environment, and thereby indirectly in the embedded molecule, but assumes a decoupled description of the embedded molecule–environment vibrational modes. A compromise can be obtained by extending the quantum region.

Before moving to complex environments such as proteins, it is important to test the performance of the PE model with respect to vibrational properties by considering more simple cases such as solvents. We here consider the vibrational Stark shift of the carbonyl (C═O) stretching mode of acetophenone in various solvents. Acetophenone is a model compound for the unnatural amino acid *p*-acetylphenylalanine, which can be site-specifically introduced into proteins by several methods.[29] The carbonyl probe has attracted increasing attention due to the observations that (i) it has a rather large oscillator strength and Stark tuning rate,[30] (ii) its frequency varies linearly with the field also for hydrogen bonding solvents, as opposed to nitrile vibrations,[31] thus making it applicable across both hydrogen and non-hydrogen bonding environments, and (iii) its experimental use in proteins has been demonstrated by overcoming the complication related to the C═O spectral overlap with the intrinsic amide I vibration of the peptide bonds[32] by a suitable choice of reference sample.[30,33] Here, we perform classical MD simulations of the acetophenone–solvent systems, and the PE–DFT model is used to compute the C═O vibrational frequency as well as the corresponding local electric field sensed by acetophenone in the geometrically frozen solvent cages for each MD configuration.

The remainder of this paper is organized in the following way. The PE model is outlined in Sec. II A and the associated molecular gradient contributions are presented in Sec. II B. Computational details are given in Sec. III. The results for acetophenone in various solvents obtained using the presented extension of the PE model are discussed in Sec. IV. Finally, concluding remarks are given in Sec. V.

II. THEORY

The key feature of the PE model is the inclusion of the electronic response of the environment to the density of the quantum region. Polarization is modeled by distributed anisotropic electronic dipole–dipole polarizabilities assigned to the atomic sites in the environment. The electrostatic energy is described by distributed permanent multipole moments. In the current version of the PE model, the effects of dispersion and exchange–repulsion interactions between the quantum region and the environment enter collectively only at the energy level through the use of Lennard-Jones (LJ) 12–6 potentials.

Our procedure for computing vibrational frequencies within the PE scheme is based on statistically independent configurations taken from an MD simulation, on which we perform a partial geometry optimization of the quantum region using the PE approach, keeping the environment fixed. The vibrational frequencies are then computed via a partial Hessian analysis of the optimized quantum region embedded in the polarizable environment (i.e., the geometry of the environment is kept fixed), where the numerical Hessian is computed based on analytic gradients. We here derive the PE contributions to the molecular gradients with respect to displacements of the nuclei in the quantum region within a variational ansatz for the quantum region, covering HF, DFT, and MCSCF levels of theories. The analytic first-order geometric derivatives of the PE energy have been implemented in a development version of the PE library[34,35] which has been interfaced to a development version of the Dalton program.[36] The required one-electron integrals are computed using the general one-electron integral framework Gen1Int by Gao, Thorvaldsen, and Ruud.[37,38] We note that molecular gradients in combination with polarizable and explicit environments have previously been reported in Refs. 39–41.

In Subsection II A, we give a brief outline of the additional energy contributions introduced by the polarizable environment and proceed to describe our implementation of the first-order geometric derivatives of the PE energy terms.

A. The polarizable embedding energy

The free energy functional for the embedded molecule can be partitioned into two contributions

$$\mathcal{E}(\boldsymbol{X}) = \mathcal{E}_{\mathrm{vac}}(\boldsymbol{X}) + \mathcal{E}_{\mathrm{PE}}(\boldsymbol{X}), \tag{2}$$

where $\mathcal{E}_{\mathrm{vac}}$ is the energy of the isolated quantum region (including wave function polarization) and $\mathcal{E}_{\mathrm{PE}}$ is the energy

originating from the interaction between the quantum region and the polarizable environment. We emphasize that the implementation covers all variational models in Dalton, not only DFT but also MCSCF, which can be used to describe molecules where static correlation is important. The energy is written as an explicit function of the vector $\boldsymbol{X}$ of the nuclear coordinates of the quantum region, for reasons that will become apparent later. The PE energy can be decomposed as

$$\mathcal{E}_{\mathrm{PE}}(\boldsymbol{X}) = \mathcal{E}_{\mathrm{es}}(\boldsymbol{X}) + \mathcal{E}_{\mathrm{ind}}(\boldsymbol{X}) + \mathcal{E}_{\mathrm{LJ}}(\boldsymbol{X}), \tag{3}$$

where $\mathcal{E}_{\mathrm{es}}$ and $\mathcal{E}_{\mathrm{ind}}$ are electrostatic contributions arising from interactions between the quantum region and the permanent and induced charge distribution of the environment, respectively, and $\mathcal{E}_{\mathrm{LJ}}$ is the LJ energy.

The electrostatic contribution $\mathcal{E}_{\mathrm{es}}$ accounts for the interactions between the permanent multipoles in the environment and the electrons and nuclei in the quantum region. In a multi-index notation,[42] these terms can be written as

$$\begin{aligned}\mathcal{E}_{\mathrm{es}}(\boldsymbol{X}) &= \mathcal{E}_{\mathrm{mul,e}}(\boldsymbol{X}) + \mathcal{E}_{\mathrm{mul,n}}(\boldsymbol{X}) \\ &= \sum_{s=1}^{S} \sum_{|k|=0}^{K_s} \frac{(-1)^{|k|}}{k!} M_s^{(k)} \\ &\quad \times \left(\sum_{pq \in \mathrm{QM}} t_{pq,s}^{(k)} D_{pq} + \sum_{m \in \mathrm{QM}} T_{sm}^{(k)} Z_m \right),\end{aligned} \tag{4}$$

where D_{pq} is an element of the one-electron density matrix in the molecular orbital (MO) basis $\{\phi_p\}$. The norm and factorial of the multi-index $k = (k_x, k_y, k_z)$ are defined as $|k| = k_x + k_y + k_z$ and $k! = k_x! k_y! k_z!$, respectively. The summation over the $|k|$ norm in Eq. (4) is over the $(|k|+1)(|k|+2)/2$ multi-indices for a given $|k|$ (e.g., for $|k| = 1$, the summation is over the three components (1,0,0), (0,1,0), and (0,0,1)). For later use, note that addition/subtraction of two multi-indices proceeds component-wise, i.e., $k \pm l = (k_x \pm l_x, k_y \pm l_y, k_z \pm l_z)$. $M_s^{(k)}$ is a Cartesian component of the $|k|$'th-order multipole moment located at the s'th site of the environment, and Z_m is the nuclear charge of m'th nucleus in the quantum region. K_s is the truncation level of the multipole expansion assigned to the s'th site. A component of the $|k|$'th-order interaction tensor is defined as a $|k|$'th-order derivative of the potential

$$T_{ij}^{(k)} = \frac{\partial^{k_x+k_y+k_z}}{\partial x_j^{k_x} \partial y_j^{k_y} \partial z_j^{k_z}} \left(\frac{1}{|\boldsymbol{r}_j - \boldsymbol{r}_i|} \right), \tag{5}$$

where i and j here denote two arbitrary sites. The $|k|$'th-order interaction integrals are then given as

$$t_{pq,s}^{(k)} = -\langle \phi_p | T_{si}^{(k)} | \phi_q \rangle, \tag{6}$$

where i refers to an electronic coordinate.

The induction energy due to the polarization of the environment both internally and by the electrons and nuclei of the quantum region is half the corresponding interaction energy

$$\mathcal{E}_{\mathrm{ind}}(\boldsymbol{X}) = -\frac{1}{2} \bar{\boldsymbol{\mu}}^T \mathbf{F} = -\frac{1}{2} \bar{\boldsymbol{\mu}}^T (\mathbf{F}_{\mathrm{mul}} + \langle \hat{\mathbf{F}}_{\mathrm{e}} \rangle + \mathbf{F}_{\mathrm{n}}). \tag{7}$$

Here, $\bar{\boldsymbol{\mu}}$ is the $3S$-dimensional column vector of the induced dipole moments at the sites in the environment, whereas $\mathbf{F}_{\mathrm{mul}}$, $\langle \hat{\mathbf{F}}_{\mathrm{e}} \rangle$, and $\mathbf{F}_{\mathrm{n}}$ are the electric fields produced at the positions of the induced dipoles by the permanent multipoles in the environment and by the electrons and nuclei in the quantum region, respectively. Superscript T denotes transposition.

The induced dipoles are determined such that the total energy is stationary, which corresponds to solving the following set of linear equations:

$$\bar{\boldsymbol{\mu}} = \boldsymbol{B}(\mathbf{F}_{\mathrm{mul}} + \langle \hat{\mathbf{F}}_{\mathrm{e}} \rangle + \mathbf{F}_{\mathrm{n}}), \tag{8}$$

with $\boldsymbol{B}$ being the $(3S \times 3S)$-dimensional classical response matrix

$$\boldsymbol{B} = \begin{pmatrix} \boldsymbol{\alpha}_1^{-1} & -\boldsymbol{T}_{12}^{(2)} & \cdots & -\boldsymbol{T}_{1S}^{(2)} \\ -\boldsymbol{T}_{21}^{(2)} & \boldsymbol{\alpha}_2^{-1} & \ddots & \vdots \\ \vdots & \ddots & \ddots & -\boldsymbol{T}_{(S-1)S}^{(2)} \\ -\boldsymbol{T}_{S1}^{(2)} & \cdots & -\boldsymbol{T}_{S(S-1)}^{(2)} & \boldsymbol{\alpha}_S^{-1} \end{pmatrix}^{-1}, \tag{9}$$

which contains the inverse of the distributed electronic dipole–dipole polarizability tensors on the diagonal and second-order interaction tensors in the off-diagonal blocks. Boldfacing is used to emphasize the tensor nature of these quantities. $\boldsymbol{B}$ is symmetric due to the polarizability tensors employed in the PE scheme.

Finally, the non-classical part of the interaction energy $\mathcal{E}_{\mathrm{LJ}}$ is, in this work, modeled by a LJ 12–6 potential

$$\begin{aligned}\mathcal{E}_{\mathrm{LJ}}(\boldsymbol{X}) &= \sum_{s=1}^{S} \sum_{m \in \mathrm{QM}} \varepsilon_{sm} \\ &\quad \times \left[\left(\frac{r_{sm}}{|\boldsymbol{R}_m - \boldsymbol{R}_s|} \right)^{12} - 2 \left(\frac{r_{sm}}{|\boldsymbol{R}_m - \boldsymbol{R}_s|} \right)^{6} \right],\end{aligned} \tag{10}$$

where the equilibrium bond length r_{sm} and well depth ε_{sm} are obtained using the Lorentz–Berthelot combination rules, i.e., $r_{sm} = \frac{1}{2}(r_s + r_m)$ and $\varepsilon_{sm} = \sqrt{\varepsilon_s \varepsilon_m}$, where r_s and ε_s are the atomic radius and well depth, respectively, of atom s and likewise for atom m.

B. Derivatives of the PE energy terms

We now turn our attention to the derivatives of the free energy in Eq. (2) with respect to a nuclear coordinate a belonging to the quantum region. The implementation of the PE contributions to the molecular gradient follows the second quantization formalism given in Ref. 43.

By straightforward differentiation and use of Wigner's $2n+1$ rule,[44] the first-order total derivatives of the free energy with respect to quantum-region nuclear displacements and evaluated at the reference geometry $\boldsymbol{X}_0$ can be written as[43,45]

$$\mathcal{E}^a(\boldsymbol{X}_0) = \left. \frac{\partial \mathcal{E}(\boldsymbol{X})}{\partial \boldsymbol{X}_a} \right|_{\boldsymbol{X}_0} = \boldsymbol{E}^a. \tag{11}$$

The explicit differentiation on the r.h.s. of Eq. (11) (denoted with superscript a) cover the Hellmann-Feynman terms as well as the integral and Pulay forces. By noting that the environment enters the energy functional of the quantum region as zero- and one-electron contributions only, the PE part of Eq. (11) can simply be added to the corresponding zero- and one-electron vacuum contributions and treated as detailed in Ref. 43, where the handling of the Pulay forces follows

implicitly from the PE contributions to the Fock matrix. It therefore suffices to consider the evaluation of the direct PE contributions to the first-order geometric derivatives of the energy.

Partial differentiation of the PE energy in Eq. (3) gives

$$\boldsymbol{E}^a_{\mathrm{PE}} = \boldsymbol{E}^a_{\mathrm{es}} + \boldsymbol{E}^a_{\mathrm{ind}} + \boldsymbol{E}^a_{\mathrm{LJ}}. \tag{12}$$

We begin by considering the electrostatic part which, according to Eq. (4), can be partitioned as

$$\boldsymbol{E}^a_{\mathrm{es}} = \boldsymbol{E}^a_{\mathrm{mul,e}} + \boldsymbol{E}^a_{\mathrm{mul,n}}. \tag{13}$$

A component a_l of the nuclear part becomes

$$E^{a_l}_{\mathrm{mul,n}} = \sum_{s=1}^{S}\left(\sum_{|k|=0}^{K_s}\frac{(-1)^{|k|}}{k!}M^{(k)}_s T^{(k+l)}_{sa} Z_a\right), \tag{14}$$

where the multi-index l denotes a Cartesian component of a quantum-region nuclear coordinate, which in the case of the gradient is truncated at first order, i.e., $|l| = 1$.

The electronic contribution is given by

$$\boldsymbol{E}^a_{\mathrm{mul,e}} = \sum_{s=1}^{S}\sum_{|k|=0}^{K_s}\frac{(-1)^{|k|}}{k!}M^{(k)}_s \sum_{pq\in\mathrm{QM}} \boldsymbol{t}^{(k)a}_{pq,s} D_{pq}, \tag{15}$$

where all three components for a given nucleus a are considered simultaneously. As the operator in Eq. (15) is independent of the nuclear coordinates belonging to the quantum region, the derivatives of the integrals involve only differentiation of the orbitals

$$\boldsymbol{t}^{(k)a}_{pq,s} = -\left(\left\langle \phi^a_p \middle| T^{(k)}_{si} \middle| \phi_q \right\rangle + \left\langle \phi_p \middle| T^{(k)}_{si} \middle| \phi^a_q \right\rangle\right). \tag{16}$$

We now proceed to the contribution from the induction energy. Recalling that $\mathbf{F}_{\mathrm{mul}}$ and $\boldsymbol{B}$ are purely PE-related quantities and $\boldsymbol{B}$ is symmetric, we obtain

$$\boldsymbol{E}^a_{\mathrm{ind}} = -\bar{\boldsymbol{\mu}}^T(\mathbf{F}^a_{\mathrm{n}} + \mathbf{F}^a_{\mathrm{e}}), \tag{17}$$

where the derivative of the electric fields generated by the nuclei in the quantum region at site s is

$$\mathbf{F}^a_{\mathrm{n},s} = -\sum_{m=1}^{M} Z_m \boldsymbol{T}^{(1)a}_{ms} = Z_a \boldsymbol{T}^{(2)}_{as}. \tag{18}$$

The geometric derivative of the corresponding field produced by the electrons in the quantum region is given by

$$\mathbf{F}^a_{\mathrm{e},s} = -\sum_{pq\in\mathrm{QM}} \boldsymbol{t}^{(1)a}_{pq,s} D_{pq}, \tag{19}$$

with the derivative integrals defined in Eq. (16).

Finally, differentiation of the LJ energy with respect to the displacement of nucleus a in the quantum region yields

$$\boldsymbol{E}^a_{\mathrm{LJ}} = -12\sum_{s=1}^{S}\varepsilon_{sa}\frac{(\boldsymbol{R}_a - \boldsymbol{R}_s)}{|\boldsymbol{R}_a - \boldsymbol{R}_s|^2} \times\left\{\left(\frac{r_{sa}}{|\boldsymbol{R}_a - \boldsymbol{R}_s|}\right)^{12} - \left(\frac{r_{sa}}{|\boldsymbol{R}_a - \boldsymbol{R}_s|}\right)^{6}\right\}. \tag{20}$$

The analytical gradients have been checked against numerical finite difference results and discrepancies were found to be below the numerical accuracy.

III. COMPUTATIONAL DETAILS

A. Preparation of the structures

The sampling of structures of acetophenone in different solvents was performed using classical MD simulations. Starting configurations were made by putting acetophenone in a cubic box with edges of 55 Å and filling the rest of the box with solvent molecules. Seven different solvents were considered: the resulting starting structures contained one acetophenone molecule with either 5430 water, 1272 dimethyl sulfoxide (DMSO), 1137 dichloromethane (DCM), 1026 chloroform, 1008 tetrahydrofuran (THF), 848 diethyl ether (DEE), or 642 *n*-hexane molecules. The systems were minimized and equilibrated in GROMACS version 4.6.3[46] using three-dimensional periodic boundary conditions. The simple point-charge (SPC) model[47] was used for water and the optimized potentials for liquid simulations (OPLS) force field[48] for all other molecules. The topologies for DCM, THF, DMSO, and chloroform were taken from the GROMACS Molecule & Liquid Database.[49,50] Non-bonded interactions were treated with a cutoff radius of 15 Å. Electrostatic interactions beyond this threshold were treated with the smooth particle-mesh Ewald method[51] with a tolerance of 10^{-5}. The minimization consisted of 20 steps of steepest descent followed by 1000 steps of conjugate gradient. To optimize the density of the system, an NPT equilibration with time step 1 fs and length 500 ps was performed using the Berendsen[52] temperature (298 K) and pressure (1 bar) coupling. Initial velocities were obtained from a Maxwell distribution at 298 K. The lengths of the cubic simulation boxes were between 51 and 55 Å for all simulation boxes after the NPT equilibration. Subsequently, the systems were equilibrated for 2 ns in the NVT ensemble using the Berendsen thermostat[52] at 298 K and a time step of 1 fs. The systems were simulated for 40 ns in the NVT ensemble with the same time step, non-bonded interactions, and other parameters as in the equilibration step. A snapshot was made every 1 ns, generating 40 uncorrelated snapshots in total. The largest fluctuations in the C═O stretching mode occur in water (see Sec. IV). Thus, to verify that the sampling was sufficiently complete, the mean and standard deviation of the C═O harmonic vibrational frequency were compared to the corresponding values based on 200 snapshots sampled in water at intervals of 200 ps. The average frequency and the associated standard deviation calculated over 200 snapshots deviated by only 0.1 cm^{-1} and 0.3 cm^{-1}, respectively, from the average over 40 snapshots, so results based on 40 snapshots were used throughout this work.

The structures used for the PE–DFT calculations are spherical systems centered around acetophenone including all solvent molecules with at least one atom within 15 Å from the solute. A thorough analysis of electrostatic and polarization interactions has shown that 15 Å is sufficient to capture the effect of a surrounding homogeneous medium such as a solvent.[53]

B. Generation of embedding potentials

The solvent embedding potentials for the PE–DFT calculations consist of quantum mechanics (QM)-derived permanent electric multipole moments up to and including quadrupoles and anisotropic electronic dipole–dipole polarizabilities located at the atomic centers as well as AMBER[54] LJ

parameters for all solvent atoms. For each snapshot, multipole moments and polarizabilities were computed explicitly for the individual solvent molecules according to the LoProp[55] approach implemented in MOLCAS,[56,57] employing the B3LYP[58–60] exchange–correlation functional and the aug-cc-pVDZ[61] basis set. The generation of the embedding potentials was facilitated by the PE Assistant Script (PEAS).[35] The basis set was recontracted to an atomic natural orbital type basis as required for the LoProp approach.[55] Additional embedding potentials (denoted PC) were made using point charges and AMBER LJ parameters. The point charges were generated in accordance with the general AMBER force field (GAFF) protocol[62] by calculating HF/6-31G* charges for every atom type using the restrained electrostatic potential (RESP) approach.[63] However, in this case, the final parameters were constructed as averages over 1000 different solvent conformations.

C. Frequency calculations

PE– and PC–B3LYP/cc-pVDZ[61] geometry optimizations were performed on each MD snapshot, optimizing acetophenone in the presence of the embedding potential representing the solvent, while fixing the solvent molecules at the configuration obtained from the MD trajectory. The Hessian of the quantum region was then computed numerically from the analytical gradients using displacements of 0.01 a.u. A reduced threshold of 10^{-4} a.u. was used for the geometries and Hessians to speed-up the calculations without sacrificing the accuracy. Subsequently, harmonic vibrational frequencies of acetophenone in the seven solvents were computed by diagonalizing the Hessian matrix for the quantum region in the fixed environment, also known as a partial normal mode analysis.[64] The C═O stretching mode is a rather localized mode, and thus, it can be reasonably well described within the vibrationally decoupled solute–solvent picture provided by a partial Hessian analysis.[65]

We note that a basis set of triple-ζ quality is generally recommended for geometry optimizations and harmonic vibrational frequencies when using the B3LYP exchange–correlation functional.[66] However, our focus is on vibrational frequency shifts rather than on absolute frequencies, where the basis set limitation is expected to be of less importance. The LJ parameters for the quantum region were taken from Ref. 67, where the parameters have been re-fitted for use in a hybrid quantum mechanics/molecular mechanics (QM/MM) scheme at the B3LYP(6-31G*)/AMBER/TIP3P level of theory. The calculations were performed using a development version of the Dalton program.[36,68] For comparison, harmonic vibrational frequencies were also computed with the polarizable continuum model[69,70] (PCM) at the same level of theory. In these cases, the solvents were described in terms of their static dielectric constants, i.e., water: 78.3553, DMSO: 46.826, DCM: 8.93, chloroform: 4.7113, THF: 7.4257, DEE: 4.24, and *n*-hexane: 1.8819. The radii of the interlocking spheres were taken from the universal force field and scaled by a factor of 1.1. The PCM calculations were carried out in Gaussian 09.[71]

To correlate electric fields and vibrational frequencies, the projections of the electric fields produced by the solvents, $\mathbf{F}_{C=O} = F_{C=O}\hat{\mathbf{u}}_{C=O}$, onto the C═O bond unit vector ($\hat{\mathbf{u}}_{C=O}$) at the center-of-mass (COM) of the C═O group of acetophenone were computed for each geometry-optimized solute–solvent configuration and include the contributions from both the permanent multipole moments and the self-consistent induced dipole moments. In this paper, we use the physics convention for the direction of the dipole moment, i.e., the dipole vector points from the negative to the positive pole, whereas the electric field is defined in the opposite manner, pointing from the positive to the negative source. Thus, a negative electric field projection onto the C═O bond vector corresponds to a stabilization of the dipole by the electric field. Moreover, we computed the electric field drop along the C═O bond defined by subtracting the field projection evaluated at the C-atom from that evaluated at the O-atom.

IV. RESULTS AND DISCUSSION

In this section, we discuss the results of the PE–DFT calculations of the vibrational Stark shifts of acetophenone in various solvents and compare to those obtained using the PC electrostatic embedding potential.

Figures 1(a) and 1(b) show the relation between the ensemble-averaged C═O harmonic vibrational frequencies of acetophenone and the associated scalar total electric field projections along the C═O bond at its COM for acetophenone in the seven different solvents as obtained using the PE– and PC–DFT procedures, respectively. The PC embedding scheme, which includes solvent polarization only in an implicit fashion via overestimated charges, is often used in QM/MM geometry optimizations and property calculations and is thus interesting from a comparative point of view. The results are also tabulated in Table I. The slope of a linear fit of the PE–DFT data (cf. Eq. (1)) yields a Stark tuning rate of 0.34 $cm^{-1}/(MV\ cm^{-1})$ ($R^2 = 0.95$). This value compares well with the microscopic tuning rates of 0.414 $cm^{-1}/(MV\ cm^{-1})$ ($R^2 = 0.99$)[30] and 0.484 $cm^{-1}/(MV\ cm^{-1})$ ($R^2 = 0.98$),[72] derived from a linear correlation of experimental frequencies and electric field projections onto the C═O bond based on MD runs with a GAFF point-charge force field and the AMOEBA polarizable force field, respectively. On the other hand, the Stark tuning rate of 0.15 $cm^{-1}/(MV\ cm^{-1})$ computed within the PC–DFT scheme is significantly underestimated. We note that the microscopic tuning rates depend on the point at which the electric field has been evaluated. The microscopic tuning rates reported by Fried *et al.* are relative to the C═O bond midpoint, whereas we used the COM of the C═O moiety. This difference will affect the case of acetophenone in water, where a significant field gradient exists along the bond (see Table I). Evaluating the field at the bond midpoint in our case would yield a lower field strength on average and thus a somewhat larger Stark tuning rate. A notable difference to the tuning rates reported by Fried *et al.* is that the one predicted in this study is based on computed quantities only, including the harmonic vibrational frequencies.

A. Electric fields

The average electric fields in Table I derived using PE–DFT are similar to the GAFF-based results in Ref. 30, apart

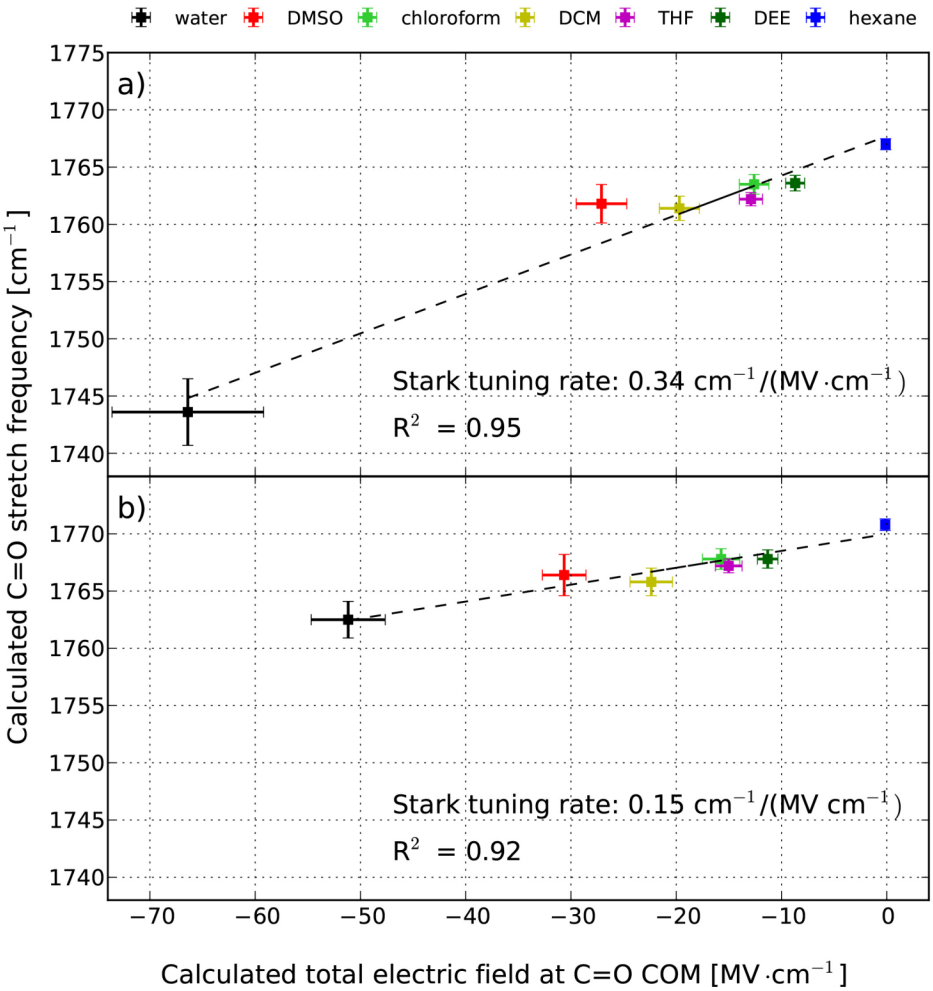

FIG. 1. C=O stretching frequency (in cm^{-1}) and total electric field projection (in MV cm^{-1}) along the C=O bond at the center-of-mass for acetophenone in different solvents. The frequencies and the electric fields were calculated with (a) PE–DFT (B3LYP/cc-pVDZ) using QM-derived distributed permanent multipoles up to quadrupole moments and distributed electronic dipole–dipole polarizabilities on all solvent atoms and (b) PC–DFT (B3LYP/cc-pVDZ) using averaged HF/6-31G*-charges on all solvent atoms. The error bars are standard errors from sampling ($N = 40$). The best-linear fit is the dashed line with the slope representing the microscopic Stark tuning rate.

from a smaller field produced by chloroform. The PC–DFT embedding scheme gives a trend for the frequency/field relation similar to the PE–DFT counterpart with the exception of water, where the PC gives a significantly smaller field strength. From a decomposition of the total PE electric field projections into contributions arising from the permanent multipole moments and induced dipoles, we find that the field from the permanent multipoles dominates the total field in all the solvents considered (where the total field was calculated as the sum of the absolute values of the field projections arising from the permanent multipole moments and induced dipoles, respectively). The induced field constitutes 42% of the total field for hexane, 33% for water, 25% for DMSO, and less than 20% for the other solvents. The magnitudes of the PE fields are, however, quantitatively very different from the fields computed using the AMOEBA force field.[72] The electrostatic components of the embedding potentials for solvent molecules used in the PE–DFT calculations were recently benchmarked against full QM calculations and were found to provide high-quality electrostatic potentials.[73] Taken together with the similarities to the AMOEBA protocol (both with respect to the electrostatic model and the level of theory) for generating the electrostatic parameters, the PE/AMOEBA field discrepancy seems to not be related to the electrostatic portion of the interaction used in

TABLE I. Ensemble-averaged C=O stretching frequencies (ν_X in cm^{-1}, where X = PE, PC) and scalar total electric field projections ($F^X_{C=O}$ in MV cm^{-1}) along the C=O bond of acetophenone in seven solvents, probed at the center-of-mass of the C=O bond. Results obtained using both the PE– and PC–B3LYP/cc-pVDZ embedding schemes are given. The induced part of the electric field projections ($F^{ind}_{C=O}$) and the field drop across the carbonyl bond ($\Delta F^{PE}_{C=O}$) are shown for the PE–DFT calculations. Standard errors from MD sampling (N = 40) are shown in parentheses. PCM frequencies (ν_{PCM}) are included for comparison.

Solvent	ν_{PE}	ν_{PC}	ν_{PCM}	ν_{exp}[a]	$F^{PE}_{C=O}$	$\Delta F^{PE}_{C=O}$	$F^{ind}_{C=O}$	$F^{PC}_{C=O}$
Water	1743.6 (2.9)	1762.5 (1.6)	1742.3	1669.4	−66.4 (7.2)	−50.3 (5.8)	−6.6 (5.9)	−51.2 (3.5)
DMSO	1761.8 (1.7)	1766.4 (1.8)	1742.8	1682.0	−27.1 (2.4)	−3.8 (1.8)	−0.6 (1.5)	−30.6 (2.1)
Chloroform	1763.5 (0.9)	1767.8 (0.9)	1751.0	1683.3	−12.6 (1.4)	−5.5 (1.1)	0.3 (0.5)	−15.7 (1.8)
DCM	1761.4 (1.1)	1765.8 (1.2)	1747.0	1684.6	−19.7 (1.9)	−5.7 (1.2)	−0.4 (0.5)	−22.4 (2.0)
THF	1762.2 (0.6)	1767.2 (0.6)	1747.9	1689.5	−12.9 (1.1)	−1.0 (1.0)	0.4 (0.5)	−15.0 (1.3)
DEE	1763.6 (0.7)	1767.8 (0.8)	1751.9	1692.9	−8.7 (0.9)	−0.7 (0.6)	0.7 (0.3)	−11.3 (1.0)
Hexane	1767.0 (0.5)	1770.8 (0.5)	1760.8	1696.4	−0.1 (0.4)	−0.7 (0.4)	0.0 (0.1)	0.1 (0.2)

[a]Reference 30.

the PE model. Following the discussion in Ref. 72, the discrepancy can rather be traced to the solute–solvent configurations produced by the point-charge MD run employed in the present study (see below).

B. Solute–solvent configurations

Radial distribution functions (RDFs) from our non-polarizable MD simulations of the solvent molecules around the carbonyl oxygen of acetophenone show that water is the only solvent with clear H-bonding character. In fact, the first peak of the RDF is just below 3 Å for water and just above 4 Å for the other six solvents, while the closest contact is at 2.4 Å for water and between 2.7 and 3.1 Å for the other six solvents. Thus, in contrast to Ref. 30, we find no evidence for H-bonding with chloroform from the present MD runs. The H-bonding in water can explain why the variation in C═O frequency is much larger in water than in the other solvents.[30] Indeed, the carbonyl oxygen atom can participate in zero, one, or two H-bonds in the different snapshots, which has a rather large impact on the electrostatic field experienced by the C═O bond. The proximity of the water solvent molecules to acetophenone is also reflected in the large field drop along the C═O bond (Table I). This large field drop is not found for the other solvents, in contrast to results from Fried *et al.* based on calculated fields from the AMOEBA polarizable MD trajectories,[72] where field drops over 15 MV cm^{-1} were also found for THF and DMSO. From the comparison with the polarizable MD runs in Ref. 72, it is clear that the lack of favorable induction interactions in the MD simulations produces too "loose" solute–solvent configurations. This poor starting point cannot be remedied in the subsequent PE–DFT geometry optimization on acetophenone with fixed solvent molecules, and thus, it gives rise to underestimated electric fields and field drops.

C. Performance of the embedding potentials

To examine the quality of the full embedding potentials in a setting that excludes the impact of the MD run, we computed one-dimensional potential energy curves along the C═O bond of acetophenone in increments of 0.01 Å. The results were compared to those obtained using an enlarged quantum region in the PE–DFT calculation, including solvent molecules within 4.0 Å of acetophenone. The results for a single and arbitrarily chosen MD configuration in water and DMSO are depicted in Figures 2(a) and 2(b), respectively. We also tested the performances of the PC and PE-nopol schemes, where the latter is the electrostatic embedding potential obtained by turning off polarization in the PE potential.

For water, we find an excellent agreement between the PE potential curves generated on the basis of the small and

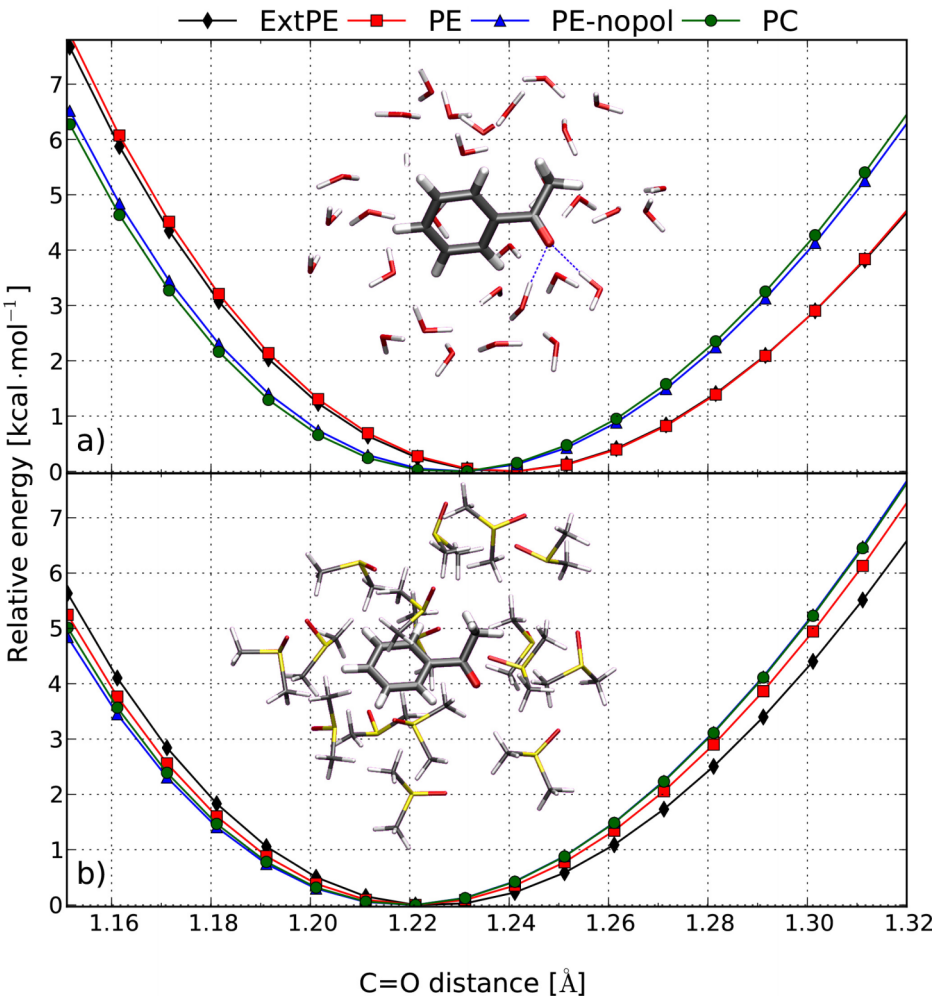

FIG. 2. C═O stretching potential energy curve for an arbitrarily chosen (a) acetophenone–water and (b) acetophenone–DMSO configuration. All atoms except oxygen are kept fixed. Single point energy calculations have been performed using PE–DFT with the following quantum regions: (i) acetophenone and solvent molecules within 4 Å (water: 35, DMSO: 16) as shown in the insets (denoted ExtPE) and (ii) acetophenone only (denoted PE). Comparison to the two electrostatic embedding schemes: (i) PC–DFT calculations employing averaged HF/6-31G* charges representing all the solvent molecules (PC model) and (ii) PE-nopol–DFT, which contains the same permanent multipole moments as PE but no polarizabilities. The quantum regions were in all cases treated at the B3LYP/cc-pVDZ level of theory.

extended quantum regions. In the configuration studied, acetophenone accepts two near-linear H-bonds from neighboring water molecules (cf. inset of Figure 2(a)). This indicates that the PE potential captures the major part of the specific interactions between acetophenone and water. In other words, the main contribution is seemingly of electrostatic origin, whereas quantum effects such as overlap and charge-transfer are less important in this case. This is in line with the prevalence of the linear Stark effect for the C═O stretch mode of acetophenone across both H-bonding and non H-bonding solvents.[31] From an inspection of the various energy terms to the interaction (at the minimum) between the quantum and classical regions, we find the electrostatic energy dominates (60%), whereas induction and exchange–repulsion/dispersion each contribute about 20%. The relative importance of electrostatics and induction to the interaction energy is thus in agreement with the relative magnitudes of the corresponding contributions to the field projection. For the energy derivatives, however, the impact of polarization becomes more pronounced, as can be seen from a comparison of the PE and PE-nopol curves. The changes in the minimum bond distance (lengthened by ~0.01 Å) and the curvature caused by the mutual coupling between the QM and PE regions are essential in order to reproduce the potential energy curve. This increase of the equilibrium distance amounts to about half the total shift induced by the solvent. Turning to the PC–DFT results, we find that the polarization, implicitly included in the embedding potential through prepolarized charges, is too weak (for the given LJ parameters), leading to a minimum at a too short C═O bond distance as well as a too narrow potential curve, almost coinciding with the profile obtained using PE-nopol.

We find a good alignment between the structures of acetophenone in water optimized using the PE–DFT scheme model with, respectively, the small and the extended quantum regions (keeping all water molecules fixed in the optimization). The structural differences are displayed graphically in Figure 3, where the two structures are superimposed. The corresponding geometric parameters are reported in Table II together with those derived using the PC–DFT model. It is clear that an overall improvement is gained by allowing the water molecules to become polarized in the PE scheme, particularly for the C═O bond length.

For the acetophenone–DMSO configuration, on the other hand, the PE–, PE-nopol–, and PC–DFT curves are quite similar with the small variations being ascribed to the lack of induction in the PC and PE-nopol potentials. The interaction is now dictated by dispersion and exchange–repulsion (70%), while electrostatics and induction only account for 25% and 5%, respectively. Indeed, the effect of dispersion and exchange–repulsion increases steadily with reduced polarity of the solvents. The difference between the PE potential curves using small and large quantum regions can thus be attributed to an inadequate description of the quantum effects in the interaction through the LJ potentials. This leads to slightly underestimated bond lengths in both the PE and PC embedding schemes.

In short, the geometrical changes induced by the polarization component of the water solvent, particularly on the C═O bond length, are significant and cannot be neglected. In apolar solvents, the influence of dispersion and exchange–repulsion increases such that the quality of the optimized structures to a larger extent depends on the employed LJ potentials.

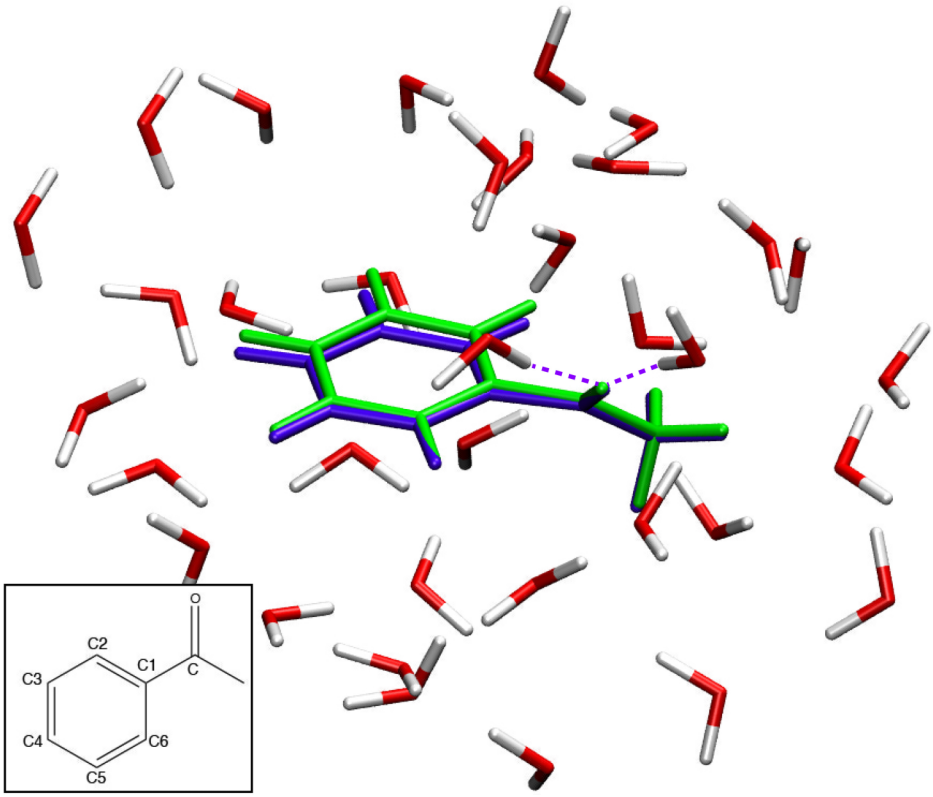

FIG. 3. Superposition of the two optimized structures for an arbitrarily chosen MD configuration of acetophenone in water obtained using PE–B3LYP/cc-pVDZ with the small (green) and extended (violet) quantum regions. All water molecules were kept fixed at the positions from the classical MD simulation. Structural parameters are listed in Table II.

TABLE II. Intra- and intermolecular structural parameters for acetophenone in water derived from PE–B3LYP/cc-pVDZ calculations with two different quantum regions: acetophenone and solvent molecules within 4 Å (ExtPE) and acetophenone only (PE) and compared to the corresponding PC–B3LYP/cc-pVDZ results. Distances are reported in Å and the dihedral angle in degrees. The atom numbering is provided in Figure 3.

	ExtPE	PE	PC
C=O	1.236	1.235	1.224
C–C1	1.496	1.495	1.502
C1–C2	1.405	1.403	1.402
C2–C3	1.398	1.397	1.397
C3–C4	1.400	1.398	1.397
C4–C5	1.403	1.402	1.400
C5–C6	1.397	1.395	1.395
C6–C1	1.409	1.407	1.405
$\phi_{O-C-C1-C2}$	134.4	132.1	127.8
O–H_{wat1}	1.745	1.706	1.837
O–H_{wat2}	1.857	1.839	1.919

On the basis of the potential curve analysis, we briefly comment on the resulting computed vibrational frequencies. We recall that the calculated Stark tuning rate computed using the PC scheme is markedly different from the PE result (cf. Figure 1). In fact, the frequency difference between hexane and water is much lower for the PC calculations (PC: 8.3 cm^{-1}; PE: 23.4 cm^{-1}; exp: 27.0 cm^{-1}) (cf. Table I). In light of the discussion above, this can be understood from the severely underestimated electrostatic interaction in water provided by the PC model, which results in an overestimated frequency. Note also that the more compact C=O bonds (cf. Figure 2) amplify the differences in the electric field projections between the two embedding schemes. In addition to the neglect of explicit polarization, there are at least three approximations made in the PC electrostatic treatment that may be partially responsible for the differences to PE: (i) the charge distribution of the solvents is represented by charges rather than by a multipole expansion up to quadrupoles; (ii) the electrostatic parameters are calculated using a lower level of theory; and (iii) the electrostatic parameters are averages over 1000 conformations rather than being calculated explicitly for every conformation separately. Judging from Figure 2, the similar potential curves provided by the PE-nopol and PC schemes indicate that the effects of (i) and (ii) are not too critical. However, to be conclusive on these points, a thorough analysis, including several solute–solvent configurations, is required. The influence of the third approximation is being investigated in ongoing work in our groups.

D. Experimental versus computed frequencies

In Table I, we have included the experimental frequencies of Fried *et al.*[30] and frequencies obtained using PCM in addition to the PE– and PC–DFT harmonic vibrational frequencies for the C=O stretch (see also Figure S1 in the supplementary material[74]). Good agreement with experimental frequencies should give a slope and correlation coefficient close to one. None of the methods show a good correlation with the experimental values. The PE–DFT frequencies show a higher slope of the linear fit (0.77) than PC–DFT (0.26) and PCM–DFT (0.58) and a correlation coefficient (0.77) similar to PC–DFT (0.84) and higher than PCM–DFT (0.66), and thus the best correlation with experiment. Discrepancies in the trend in C=O frequency across the different solvents are expected from the limitations in the computational approach: in addition to the assumption of a decoupled and harmonic C=O stretching mode, the lack of polarization in the classical MD simulation leads to imperfect geometries for both intra- and intermolecular geometric parameters that affect the Stark tuning rate. This holds in particular for the configuration of the solvent molecules whose geometries are kept fixed in the PE geometry optimization step. A more accurate way of obtaining the structures is with polarizable MD, QM/MM, or full QM MD simulations. Also related to the geometries is the rather crude modeling of exchange–repulsion and dispersion by the empirical LJ terms. Even though we used LJ parameters developed specifically for the quantum region,[67] they are optimized to be compatible with a B3LYP/6-31G*/AMBER/TIP3P embedding scheme. This leads to an imperfect balance between electrostatics, dispersion, and exchange–repulsion in the PE calculations (and apparently also for the PC calculations), as clearly seen for the acetophenone–DMSO configuration in Figure 2(b). The poor balance between LJ parameters and electrostatics in the PC scheme is also pronounced from the non-smooth transition from the apolar solvents to the computed gas phase value of 1769.8 cm^{-1}.

As an aside we note that the PCM–DFT value for the C=O frequency in water, and thereby, the PCM frequency range across the solvents, is strongly affected by the choice of PCM cavity parameters (in the simple Onsager model, the asymptotic value for the reaction field is R^{-3}, where R is the cavity radius). The good agreement between the PCM– and PE–DFT C=O frequencies in water as well as the reasonable PCM–DFT hexane to water shift does, therefore, not merit further consideration as can also be seen from the incapability of PCM to capture the trend across the solvents.

E. Effects of electronic conjugation

Returning to the vibrational Stark effect, Figure 1(a) clearly demonstrates a correlation ($R^2 = 0.95$) between the average C=O stretching frequencies and electric field projections along the C=O bond for acetophenone in various solvents. This correlation is an effect of the averaging over 40 snapshots for each of the solvents. In fact, we do not find a correlation at the level of the snapshots, i.e., when considering the 40 snapshots of each solvent as individual data points. This stands in contrast to the linearity found at the snapshot level for the C=O stretching mode of acetone in water in the work by Choi and Cho.[31] The individual frequencies, however, do correlate with the bond length of the optimized geometry of acetophenone in the PE potential of the seven solvents ($0.73 < R^2 < 0.97$) with a progressively longer C=O distance for increasingly polar solvents. Such linear frequency–bond length correlations have been demonstrated previously for numerous vibrational probes.[31,75,76] This seems to indicate that the frequency is not only determined by the electric field of the solvent but also by intramolecular factors of the probe. The lack of correlation at the snapshot level can be explained by

two competing mechanisms. The *direct field effect* is that a large electric field along the C═O bond *lengthens* the bond, thereby decreasing the C═O vibrational frequency. On the other hand, the *indirect field effect* (i.e., large field components normal to the plane defined by the acetyl group) leads, together with thermally driven fluctuations, to an out-of-plane rotation of the C═O group. The resulting non-planarity reduces the electronic resonance and is associated with a *shortening* of the C═O bond and an increased vibrational frequency. Non-linearities in the Stark tuning rate are thus introduced at the microscopic level by the varying degree of electronic conjugation. Indeed, in water and DMSO, which generate the largest electric field (Table I), the average deviation from planarity of acetophenone ($16° < \phi < 17°$) is larger than the average for the other five solvents ($6° < \phi < 11°$), where ϕ is the dihedral angle between the carbonyl group and the phenyl ring. These competing mechanisms can explain why acetophenone has a similar C═O stretching frequency in DMSO and chloroform even though the field generated by DMSO is larger (cf. Table I). Other factors that contribute to deviations from a linear correlation between frequency and field at the snapshot level are non-linearities due to the breakdown of the linear Stark effect at sufficiently high field strengths as well as the spatial field heterogeneity (cf. field drops in Table I) across the vibrational probe, which is not taken into account by the point dipole model in Eq. (1).[31] The fact that electronic conjugation plays a role is also clear from a comparison of the measured macroscopic Stark tuning rates ($|\Delta\boldsymbol{\mu}|f$, where f is the local field factor) when going from aliphatic to increasingly conjugated C═O groups; acetone: 0.8 $cm^{-1}/(MV\ cm^{-1})$[77] < acetophenone: 1.05 $cm^{-1}/(MV\ cm^{-1})$[30] < 6-propionyl-2-(dimethylamino)naphtalene: 1.8 $cm^{-1}/(MV\ cm^{-1})$.[78] In contrast to acetone,[31] the C═O group of acetophenone cannot be viewed as being electronically decoupled and the variations in electronic conjugation occur within the same molecule. However, this effect is averaged out over the snapshots and thus not manifested in the macroscopic behavior.

V. SUMMARY

In this work, we have implemented analytical molecular gradients in the PE model, which makes it possible to perform geometry optimizations and numerical Hessian analysis for molecules embedded in polarizable environments. The first application of the new implementation is on the vibrational Stark effect of the C═O stretching mode of acetophenone in seven solvents of different polarities, using solute–solvent starting configurations obtained from non-polarizable MD simulations. In particular, we focused on the impact of explicit polarization on geometries and vibrational frequencies by comparing to a typical point-charge embedding scheme.

We find that inclusion of explicit polarization in the PE scheme is essential to capture the electrostatic component of the H-bonding interaction between acetophenone and water, which induces an elongation of the C═O bond and a concomitant bathochromic frequency shift. For the more apolar solvents, where quantum effects are of increasing importance, the performance of the PE scheme is similar to that of PC embedding and limited by the description of the non-electrostatic effects modeled through LJ potentials.

The Stark tuning rate based solely on PE–DFT-computed quantities is slightly smaller than the semi-experimental values found by Fried *et al.*[30,72] It turns out that this difference largely relates to the quality of the underlying solute–solvent configurations derived from the MD simulations, which indeed is sensitive to the incorporation of polarization irrespective of the solvent, as demonstrated by Fried *et al.*[72] It is thus clear that the LJ potentials and underlying solute–solvent configurations constitute critical factors in our procedure, which require careful handling.

ACKNOWLEDGMENTS

N.H.L. and J.K. acknowledge the Lundbeck Foundation and the Danish Council for Independent Research (DFF) for financial support. M.T.P.B., B.G., and K.R. acknowledge support from the Research Council of Norway through a Centre of Excellence Grant (Grant No. 179568/V30), the European Research Council through a Starting Grant (Grant No. 279619), and the Norwegian Supercomputer Program (Grant No. NN4654K). M.T.P.B. gratefully acknowledges financial support from BioStruct, the Norwegian national graduate school in structural biology. J.M.H.O. acknowledges financial support from DFF through the Sapere Aude research career program. J.K. thanks the Villum foundation for financial support.

[1]J. Villa and A. Warshel, "Energetics and dynamics of enzymatic reactions," J. Phys. Chem. B **105**, 7887–7907 (2001).
[2]M. Strajbl, A. Shurki, M. Kato, and A. Warshel, "Apparent NAC effect in chorismate mutase reflects electrostatic transition state stabilization," J. Am. Chem. Soc. **125**, 10228–10237 (2003).
[3]A. Warshel, P. K. Sharma, M. Kato, Y. Xiang, H. Liu, and M. H. Olsson, "Electrostatic basis for enzyme catalysis," Chem. Rev. **106**, 3210–3235 (2006).
[4]S. Hammes-Schiffer, "Catalytic efficiency of enzymes: A theoretical analysis," Biochemistry **52**, 2012–2020 (2013).
[5]C. T. Liu, J. P. Layfield, R. J. Stewart, J. B. French, P. Hanoian, J. B. Asbury, S. Hammes-Schiffer, and S. J. Benkovic, "Probing the electrostatics of active site microenvironments along the catalytic cycle for *Escherichia coli* dihydrofolate reductase," J. Am. Chem. Soc. **136**, 10349–10360 (2014).
[6]M. Drobizhev, N. S. Makarov, S. E. Tillo, T. E. Hughes, and A. Rebane, "Two-photon absorption properties of fluorescent proteins," Nat. Methods **8**, 393–399 (2011).
[7]J.-Y. Hasegawa, K. J. Fujimoto, and H. Nakatsuji, "Color tuning in photofunctional proteins," ChemPhysChem **12**, 3106–3115 (2011).
[8]N. H. List, J. M. H. Olsen, H. J. A. Jensen, A. H. Steindal, and J. Kongsted, "Molecular-level insight into the spectral tuning mechanism of the DsRed chromophore," J. Phys. Chem. Lett. **3**, 3513–3521 (2012).
[9]E. S. Park, S. S. Andrews, R. B. Hu, and S. G. Boxer, "Vibrational Stark spectroscopy in proteins: A probe and calibration for electrostatic fields," J. Phys. Chem. B **103**, 9813–9817 (1999).
[10]H. Lehle, J. M. Kriegl, K. Nienhaus, P. Deng, S. Fengler, and G. U. Nienhaus, "Probing electric fields in protein cavities by using the vibrational Stark effect of carbon monoxide," Biophys. J. **88**, 1978–1990 (2005).
[11]S. G. Boxer, "Stark realities," J. Phys. Chem. B **113**, 2972–2983 (2009).
[12]B. A. Lindquist, K. E. Furse, and S. A. Corcelli, "Nitrile groups as vibrational probes of biomolecular structure and dynamics: An overview," Phys. Chem. Chem. Phys. **11**, 8119–8132 (2009).
[13]L. Xu, A. E. Cohen, and S. G. Boxer, "Electrostatic fields near the active site of human aldose reductase: 2. New inhibitors and complications caused by hydrogen bonds," Biochemistry **50**, 8311–8322 (2011).
[14]C. G. Bazewicz, J. S. Lipkin, E. E. Smith, M. T. Liskov, and S. H. Brewer, "Expanding the utility of 4-cyano-l-phenylalanine as a vibrational reporter of protein environments," J. Phys. Chem. B **116**, 10824–10831 (2012).
[15]A. T. Fafarman, P. A. Sigala, J. P. Schwans, T. D. Fenn, D. Herschlag, and S. G. Boxer, "Quantitative, directional measurement of electric field

heterogeneity in the active site of ketosteroid isomerase," Proc. Natl. Acad. Sci. U. S. A. **109**, E299–E308 (2012).
[16]H. Kim and M. Cho, "Infrared probes for studying the structure and dynamics of biomolecules," Chem. Rev. **113**, 5817–5847 (2013).
[17]P. A. Sigala, A. T. Fafarman, J. P. Schwans, S. D. Fried, T. D. Fenn, J. M. M. Caaveiro, B. Pybus, D. Ringe, G. A. Petsko, S. G. Boxer, and D. Herschlag, "Quantitative dissection of hydrogen bond-mediated proton transfer in the ketosteroid isomerase active site," Proc. Natl. Acad. Sci. U. S. A. **110**, E2552–E2561 (2013).
[18]P. R. Callis and B. K. Burgess, "Tryptophan fluorescence shifts in proteins from hybrid simulations: An electrostatic approach," J. Phys. Chem. B **101**, 9429–9432 (1997).
[19]M. Drobizhev, S. Tillo, N. S. Makarov, T. E. Hughes, and A. Rebane, "Color hues in red fluorescent proteins are due to internal quadratic Stark effect," J. Phys. Chem. B **113**, 12860–12864 (2009).
[20]P. R. Callis, "Exploring the electrostatic landscape of proteins with tryptophan fluorescence," in *Reviews in Fluorescence 2007* (Springer, 2009), pp. 199–248.
[21]O. K. Asamoah, J. P. Wuskell, L. M. Loew, and F. Bezanilla, "A fluorometric approach to local electric field measurements in a voltage-gated ion channel," Neuron **37**, 85–97 (2003).
[22]M. Drobizhev, J. N. Scott, P. R. Callis, and A. Rebane, "All-optical sensing of the components of the internal local electric field in proteins," IEEE Photonics J. **4**, 1996–2001 (2012).
[23]S. S. Andrews and S. G. Boxer, "Vibrational Stark effects of nitriles I. Methods and experimental results," J. Phys. Chem. A **104**, 11853–11863 (2000).
[24]J. Jeon, S. Yang, J.-H. Choi, and M. Cho, "Computational vibrational spectroscopy of peptides and proteins in one and two dimensions," Acc. Chem. Res. **42**, 1280–1289 (2009).
[25]J. M. Olsen, K. Aidas, and J. Kongsted, "Excited states in solution through polarizable embedding," J. Chem. Theory Comput. **6**, 3721–3734 (2010).
[26]J. M. H. Olsen and J. Kongsted, "Molecular properties through polarizable embedding," Adv. Quantum Chem. **61**, 107–143 (2011).
[27]N. H. List, F. M. Pimenta, L. Holmegaard, R. L. Jensen, M. Etzerodt, T. Schwabe, J. Kongsted, P. R. Ogilby, and O. Christiansen, "Effect of chromophore encapsulation on linear and nonlinear optical properties: The case of miniSOG, a protein-encased flavin," Phys. Chem. Chem. Phys. **16**, 9950–9959 (2014).
[28]C. Steinmann, J. M. H. Olsen, and J. Kongsted, "Nuclear magnetic shielding constants from quantum mechanical/molecular mechanical calculations using polarizable embedding: Role of the embedding potential," J. Chem. Theory Comput. **10**, 981–988 (2014).
[29]L. Wang, Z. Zhang, A. Brock, and P. G. Schultz, "Addition of the keto functional group to the genetic code of *Escherichia coli*," Proc. Natl. Acad. Sci. U. S. A. **100**, 56–61 (2003).
[30]S. D. Fried, S. Bagchi, and S. G. Boxer, "Measuring electrostatic fields in both hydrogen-bonding and non-hydrogen-bonding environments using carbonyl vibrational probes," J. Am. Chem. Soc. **135**, 11181–11192 (2013).
[31]J.-H. Choi and M. Cho, "Vibrational solvatochromism and electrochromism of infrared probe molecules containing C≡O, C≡N, C═O, or C–F vibrational chromophore," J. Chem. Phys. **134**, 154513 (2011).
[32]S. Krimm and J. Bandekar, "Vibrational spectroscopy and conformation of peptides, polypeptides, and proteins," Adv. Protein Chem. **38**, 181–364 (1986).
[33]S. D. Fried and S. G. Boxer, "Evaluation of the energetics of the concerted acid-base mechanism in enzymatic catalysis: The case of ketosteroid isomerase," J. Phys. Chem. B **116**, 690–697 (2012).
[34]J. M. H. Olsen, The Polarizable Embedding (PE) Library (Development Version), 2014.
[35]J. M. H. Olsen, "Development of quantum chemical methods towards rationalization and optimal design of photoactive proteins," Ph.D. thesis, University of Southern Denmark, Odense, Denmark, 2012.
[36]K. Aidas, C. Angeli, K. L. Bak, V. Bakken, R. Bast, L. Boman, O. Christiansen, R. Cimiraglia, S. Coriani, P. Dahle, E. K. Dalskov, U. Ekström, T. Enevoldsen, J. J. Eriksen, P. Ettenhuber, B. Fernández, L. Ferrighi, H. Fliegl, L. Frediani, K. Hald, A. Halkier, C. Hättig, H. Heiberg, T. Helgaker, A. C. Hennum, H. Hettema, E. Hjertenæs, S. Høst, I.-M. Høyvik, M. F. Iozzi, B. Jansik, H. J. A. Jensen, D. Jonsson, P. Jørgensen, J. Kauczor, S. Kirpekar, T. Kjærgaard, W. Klopper, S. Knecht, R. Kobayashi, H. Koch, J. Kongsted, A. Krapp, K. Kristensen, A. Ligabue, O. B. Lutnæs, J. I. Melo, K. V. Mikkelsen, R. H. Myhre, C. Neiss, C. B. Nielsen, P. Norman, J. Olsen, J. M. H. Olsen, A. Osted, M. J. Packer, F. Pawlowski, T. B. Pedersen, P. F. Provasi, S. Reine, Z. Rinkevicius, T. A. Ruden, K. Ruud, V. Rybkin, P. Salek, C. C. M. Samson, A. S. de Merás, T. Saue, S. P. A. Sauer, B. Schimmelpfennig, K. Sneskov, A. H. Steindal, K. O. Sylvester-Hvid, P. R. Taylor, A. M. Teale, E. I. Tellgren, D. P. Tew, A. J. Thorvaldsen, L. Thøgersen, O. Vahtras, M. A. Watson, D. J. D. Wilson, M. Ziolkowski, and H. Ågren, "The Dalton quantum chemistry program system," WIREs Comput. Mol. Sci. **4**, 269 (2014).
[37]B. Gao, A. J. Thorvaldsen, and K. Ruud, "GEN1INT: A unified procedure for the evaluation of one-electron integrals over Gaussian basis functions and their geometric derivatives," Int. J. Quantum Chem. **111**, 858–872 (2011).
[38]B. Gao and A. J. Thorvaldsen, Gen1Int version 0.2.1, 2012, http://repo.ctcc.no/projects/gen1int.
[39]S. Caprasecca, S. Jurinovich, L. Viani, C. Curutchet, and B. Mennucci, "Geometry optimization in polarizable QM/MM models: The induced dipole formulation," J. Chem. Theory Comput. **10**(4), 1588–1598 (2014).
[40]H. Li and M. S. Gordon, "Polarization energy gradients in combined quantum mechanics, effective fragment potential, and polarizable continuum model calculations," J. Chem. Phys. **126**, 124112 (2007).
[41]J. L Payton, S. M. Morton, J. E. Moore, and L. Jensen, "A discrete interaction model/quantum mechanical method for simulating surface-enhanced Raman spectroscopy," J. Chem. Phys. **136**, 214103 (2012).
[42]X. S. Raymond, *Elementary Introduction to the Theory of Pseudodifferential Operators* (CRC Press, Boca Raton, Florida, 1991).
[43]T. U. Helgaker, J. Almlöf, H. J. A. Jensen, and P. Jørgensen, "Molecular Hessians for large-scale MCSCF wave functions," J. Chem. Phys. **84**, 6266–6279 (1986).
[44]J. G. Ángyán, "Wigners (2n + 1) rule for nonlinear Schrödinger equations," J. Math. Chem. **46**, 1–14 (2009).
[45]P. Pulay, "*Ab initio* calculation of force constants and equilibrium geometries in polyatomic molecules: I. Theory," Mol. Phys. **17**, 197 (1969).
[46]B. Hess, C. Kutzner, D. van der Spoel, and E. Lindahl, "GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation," J. Chem. Theory Comput. **4**, 435–447 (2008).
[47]H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, and J. Hermans, in *Intermolecular Forces*, edited by B. Pullman (Reidel, Dordrecht, 1981).
[48]W. L. Jorgensen, D. S. Maxwell, and J. Tirado-Rives, "Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids," J. Am. Chem. Soc. **118**, 11225–11236 (1996).
[49]C. Caleman, P. J. van Maaren, M. Hong, J. S. Hub, L. T. Costa, and D. van der Spoel, "Force field benchmark of organic liquids: Density, enthalpy of vaporization, heat capacities, surface tension, isothermal compressibility, volumetric expansion coefficient, and dielectric constant," J. Chem. Theory Comput. **8**, 61–74 (2012).
[50]D. van der Spoel, P. J. van Maaren, and C. Caleman, "GROMACS molecule & liquid database," Bioinformatics **28**, 752–753 (2012).
[51]T. Darden, D. York, and L. Pedersen, "Particle mesh Ewald: An N·log (N) method for Ewald sums in large systems," J. Chem. Phys. **98**, 10089–10092 (1993).
[52]H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak, "Molecular dynamics with coupling to an external bath," J. Chem. Phys. **81**, 3684–3690 (1984).
[53]M. T. P. Beerepoot, A. H. Steindal, K. Ruud, J. M. H. Olsen, and J. Kongsted, "Convergence of environment polarization effects in multiscale modeling of excitation energies," Comput. Theor. Chem. **1040–1041**, 304–311 (2014).
[54]W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman, "A second generation force field for the simulation of proteins, nucleic acids, and organic molecules," J. Am. Chem. Soc. **117**, 5179–5197 (1995).
[55]L. Gagliardi, R. Lindh, and G. Karlström, "Local properties of quantum chemical systems: The LoProp approach," J. Chem. Phys. **121**, 4494–4500 (2004).
[56]G. Karlström, R. Lindh, P.-Å. Malmqvist, B. O. Roos, U. Ryde, V. Veryazov, P.-O. Widmark, M. Cossi, B. Schimmelpfennig, P. Neogrady, and L. Seijo, "MOLCAS: A program package for computational chemistry," Comput. Mater. Sci. **28**, 222–239 (2003).
[57]F. Aquilante, L. De Vico, N. Ferré, G. Ghigo, P.-Å. Malmqvist, P. Neogrády, T. B. Pedersen, M. Pitoňák, M. Reiher, B. O. Roos, L. Serrano-Andrés, M. Urban, V. Veryazov, and R. Lindh, "MOLCAS 7: The next generation," J. Comput. Chem. **31**, 224–247 (2010).
[58]A. D. Becke, "Density-functional thermochemistry. III. The role of exact exchange," J. Chem. Phys. **98**, 5648–5652 (1993).
[59]S. H. Vosko, L. Wilk, and M. Nusair, "Accurate spin-dependent electron liquid correlation energies for local spin density calculations: A critical analysis," Can. J. Phys. **58**, 1200–1211 (1980).
[60]C. Lee, W. Yang, and R. G. Parr, "Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density," Phys. Rev. B **37**, 785–789 (1988).

[61]T. H. Dunning, "Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen," J. Chem. Phys. **90**, 1007–1023 (1989).
[62]J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, and D. A. Case, "Development and testing of a general amber force field," J. Comput. Chem. **25**, 1157–1174 (2004).
[63]C. I. Bayly, P. Cieplak, W. D. Cornell, and P. A. Kollman, "A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: The RESP model," J. Phys. Chem. **97**, 10269–10280 (1993).
[64]H. Li and J. H. Jensen, "Partial Hessian vibrational analysis: The localization of the molecular vibrational energy and entropy," Theor. Chem. Acc. **107**, 211–219 (2002).
[65]A. Ghysels, V. Van Speybroeck, E. Pauwels, S. Catak, B. R. Brooks, D. Van Neck, and M. Waroquier, "Comparative study of various normal mode analysis techniques based on partial Hessians," J. Comput. Chem. **31**, 994–1007 (2010).
[66]J. M. L. Martin, J. El-Yazal, and J.-P. François, "Basis set convergence and performance of density functional theory including exact exchange contributions for geometries and harmonic frequencies," Mol. Phys. **86**, 1437–1450 (1995).
[67]M. Freindorf, Y. Shao, T. R. Furlani, and J. Kong, "Lennard-Jones parameters for the combined QM/MM method using the B3LYP/6-31G*/AMBER potential," J. Comput. Chem. **26**, 1270–1278 (2005).
[68]Dalton, A molecular electronic structure program, Release DALTON2013.0, 2013, http://daltonprogram.org/.
[69]E. Cancès, B. Mennucci, and J. Tomasi, "A new integral equation formalism for the polarizable continuum model: Theoretical background and applications to isotropic and anisotropic dielectrics," J. Chem. Phys. **107**, 3032–3041 (1997).
[70]J. Tomasi, B. Mennucci, and R. Cammi, "Quantum mechanical continuum solvation models," Chem. Rev. **105**, 2999–3094 (2005).
[71]M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman , G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, GAUSSIAN 09 Revision D.01, Gaussian, Inc., Wallingford, CT, 2009.
[72]S. D. Fried, L.-P. Wang, S. G. Boxer, P. Ren, and V. S. Pande, "Calculations of the electric fields in liquid solutions," J. Phys. Chem. B **117**, 16236–16248 (2013).
[73]T. Schwabe, J. M. H. Olsen, K. Sneskov, J. Kongsted, and O. Christiansen, "Solvation effects on electronic transitions: Exploring the performance of advanced solvent potentials in polarizable embedding calculations," J. Chem. Theory Comput. **7**, 2209–2217 (2011).
[74]See supplementary material at http://dx.doi.org/10.1063/1.4905909 for correlation plots between experimental and computed C═O vibrational frequencies in different solvents.
[75]H. Lee, J.-H. Choi, and M. Cho, "Vibrational solvatochromism and electrochromism of cyanide, thiocyanate, and azide anions in water," Phys. Chem. Chem. Phys. **12**, 12658–12669 (2010).
[76]J.-H. Choi, K.-I. Oh, H. Lee, C. Lee, and M. Cho, "Nitrile and thiocyanate IR probes: Quantum chemistry calculation studies and multivariate least-square fitting analysis," J. Chem. Phys. **128**, 134506 (2008).
[77]E. S. Park and S. G. Boxer, "Origins of the sensitivity of molecular vibrations to electric fields: Carbonyl and nitrosyl stretches in model compounds and proteins," J. Phys. Chem. B **106**, 5800–5806 (2002).
[78]I. T. Suydam and S. G. Boxer, "Vibrational Stark effects calibrate the sensitivity of vibrational probes for electric fields in proteins," Biochemistry **42**, 12050–12055 (2003).