Theor Chem Acc (2007) 117:1001–1015
DOI 10.1007/s00214-006-0216-z

REGULAR ARTICLE

Theoretical modeling of spectroscopic properties of molecules in solution: toward an effective dynamical discrete/continuum approach

Giuseppe Brancato · Vincenzo Barone · Nadia Rega

Received: 13 July 2006 / Accepted: 13 October 2006 / Published online: 23 January 2007

Abstract In this work, we present an effective and flexible computational approach, which is the result of an on-going development conducted in our group, for simulating complex solute–solvent systems and computing relevant spectroscopic observables. Such an approach is based on QM/MM molecular dynamics techniques using non-periodic conductor boundary conditions and localized basis sets, combined with a posteriori high-level quantum mechanical methods for the calculation of spectroscopic parameters. As illustrative applications, we report structural and spectroscopic analyses of acetone, acrolein and glycine radical in aqueous solutions, where solvent effects on the NMR chemical shifts, UV absorption spectrum and EPR hyperfine coupling constants, respectively, are investigated and favorably compared to experimental measurements. In particular, it will be shown the importance of including dynamical effects in order to reproduce experimental data accurately. Moreover, we present an infrared analysis of formamide in both gas phase and acetonitrile from first-principle molecular dynamics simulations.

Keywords Mean field · NMR · IR · EPR · UV

1 Introduction

The theoretical modeling of solvent effects is a well-studied topic of modern theoretical chemistry, which has been extensively discussed in many specialized reviews [1–3]. Traditionally, implicit solvent models have been successfully used to provide valuable information on a large number of chemical systems in the liquid phase. Such models effectively treat complex solute–solvent systems by drastically reducing the number of degrees of freedom of the environment (solvent), which is modeled as a dielectric continuum, and, then, allowing for the use of high-level quantum mechanical methods at an affordable computational cost. Different molecular properties in solution have been reliably reproduced by means of continuum based models, such as structural changes, vibrational frequencies and also more subtle spectroscopic parameters. However, in almost all cases static calculations on a single or a few reference configurations were performed, therefore neglecting solute–solvent dynamical effects that are, in many circumstances, of fundamental importance to interpret and accurately predict experimental data.

Here, we present an effective discrete/continuum theoretical model, referred to as the mean field (MF) [4–6] model, which is particularly well-suited for treating solute–solvent systems of variable size and at different levels of theory, from purely classical force fields based methods to more sophisticated mixed quantum mechanics/molecular mechanics (QM/MM) and full quantum mechanics methodologies. The MF model was especially designed for performing effective phase–space samplings of the solute–solvent system by using molecular dynamics techniques. In the framework of our model, ab initio molecular dynamics can be performed both in the Born–Oppenheimer [7,8] and the extended-Lagrangian [9–12] formalism using localized basis functions. Note that the latter, as compared to plane wave functions, should be considered a more natural choice for the treatment of non-periodic systems and, besides, they allow to employ, in an efficient manner, well-trusted

G. Brancato · V. Barone · N. Rega (✉)
Dipartimento di Chimica and INSTM, Università Federico II, Complesso Monte S. Angelo, Via Cintia, 80126 Naples, Italy
e-mail: nadia.rega@unina.it

DFT methods based on hybrid functionals, e.g. B3LYP [13]. Also, it should be noted that although solvent bulk effects can be adequately taken into account by dielectric continuum models, specific and relatively strong interactions, such as hydrogen bonds, that may occur between a solute and the solvent are generally poorly described by such models. In some cases, solute–solvent interactions are characterized by subtle electronic quantum effects, e.g. Pauli exclusion principle and charge (or spin) transfer that necessarily require a full quantum mechanical treatment of, at least, a few solvent molecules along with the solute. Remarkably, the MF model allows to either treat implicitly the whole solvent or, if necessary, model explicitly a few solvation shells with full atomic detail.

Moreover, we will show how the combination of the MF model with a posteriori quantum mechanical calculations makes it possible to obtain a "state of the art" computational tool for the accurate evaluation of spectroscopic properties in solution. In particular, in the present work we will consider, as illustrative applications, aqueous solutions of acetone, acrolein and glycine radical, reporting a detailed comparison between computed and experimental spectroscopic parameters, such as NMR chemical shifts, UV absorption transitions and EPR hyperfine coupling constants. In all such cases, it will be shown that thermal effects have a primary role on the considered spectroscopic properties and, as a consequence, reliable theoretical results can be obtained only from statistical averages over representative solute–solvent configurational samplings. To this purpose, hybrid QM/MM molecular dynamics simulations of the above aqueous solutions have been performed at room temperature by treating the solute at full QM level with a minimum number of solvent (water) molecules, say two/three solvent shells, at a lower and computationally more favorable level (MM). Note that the solvent was included explicitly in such simulations to properly account for the directional hydrogen bonding interactions.

Furthermore, we will present a preliminary study on the dynamical behavior of formamide in acetonitrile, where the weakly interacting solvent was entirely described in terms of a structureless polarizable dielectric. Also, in this case, a comparison with experiments gave satisfactory results, even if only an indirect comparison was possible.

This paper is organized as follows. In Sect. 2, the basic theoretical formulation of the MF model is presented and its constituent terms are described in full details. Also, the computational procedure used to evaluate molecular spectroscopic parameters is discussed in Sect. 2.5 and all the required computational details are provided in Sect. 2.6. In Sect. 3, we report some structural and spectroscopic results of aqueous solutions of acetone, acrolein and glycine radical. To be specific, the solvent effects on the ^{13}C and ^{17}O NMR isotropic shielding constants of acetone are investigated and reported in Sect. 3.1, the change of UV $n \rightarrow \pi*$ transition energy of acrolein on-going from the gas phase to the aqueous solution in Sect. 3.2 and the EPR hyperfine coupling constants of glycine radical in Sect. 3.3. In Sect. 3.4, a detailed normal mode analysis of formamide in the gas phase and in acetonitrile is presented. Concluding remarks and perspectives are given in Sect. 4.

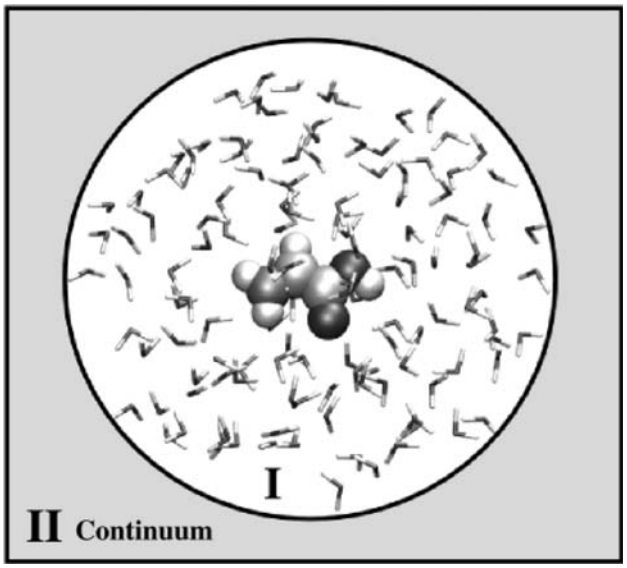

Fig. 1 Illustrative decomposition into system I (explicit system) and system II (continuum), according to the MF model

2 Methods

2.1 Basic equations

We consider a macroscopic solute–solvent system at infinite dilution in a NVT ensemble, which obviously includes the homogeneous liquid as a special case, where, say in the center, we define a microscopic subsystem, referred to as system I, which contains the solute and n solvent molecules, with $n << N$ (see Fig. 1). Hence, we can decompose the overall system in two subsystems with fixed volume and number of molecules. In other words, we can say that system I is embedded into a cavity of the macroscopic system. We assume, for the sake of simplicity, that the potential energy of the whole system can be decomposed as

$$\mathcal{U}_{\mathrm{tot}}(\boldsymbol{x}_I, \boldsymbol{x}_I I) = \mathcal{U}_I(\boldsymbol{x}_I) + \mathcal{U}_{II}(\boldsymbol{x}_{II}) + \mathcal{V}(\boldsymbol{x}_I, \boldsymbol{x}_{II}) \quad (1)$$

with $\boldsymbol{x}_I$ and $\boldsymbol{x}_{II}$ the coordinates of system I and of the remaining part of the macroscopic NVT system (II) respectively, $\mathcal{U}_I$ and $\mathcal{U}_{II}$ the internal energies and $\mathcal{V}$ the interaction energy between the two systems. Then, we have that the (Helmholtz) free energy of a given phase space position of the molecules in the cavity is

$$A(\boldsymbol{x}_I) = \mathcal{U}_I(\boldsymbol{x}_I) - kT \ln\left(\Theta^{-1} \int \mathrm{e}^{-\beta[\mathcal{V}(\boldsymbol{x}_I,\boldsymbol{x}_{II})+\mathcal{U}_{II}(\boldsymbol{x}_{II},\boldsymbol{p}_{II})]} \frac{\mathrm{d}\boldsymbol{x}_{II}\mathrm{d}\boldsymbol{p}_{II}}{h^{d_{II}}}\right) \tag{2}$$

where $\boldsymbol{x}_I, \boldsymbol{p}_I$ and $\boldsymbol{x}_{II}, \boldsymbol{p}_{II}$ are the coordinates and conjugated momenta of system I and II respectively, d_{II} is the total number of classical degrees of freedom in the macroscopic subsystem, h is the Planck's constant, Θ^{-1} is a constant providing the quantum correction and $\beta^{-1} = kT$ with k the Boltzmann's constant and T the absolute temperature. Note that $\mathcal{V}$ contains a repulsive term that prevents solvent molecules from penetrating the cavity boundaries, as described below, and consequently the integrand of Eq. 2 has zero measure in region I. If we now choose as reference state the one where all the interactions between the two subsystems are switched off ($\mathcal{V} = 0$), we can express the free energy change as

$$\begin{aligned}\Delta A(\boldsymbol{x}_I) &= A(\boldsymbol{x}_I) - A_{\mathrm{ref}}(\boldsymbol{x}_I) \\ &= -kT \ln\left(\frac{\int \mathrm{e}^{-\beta[\mathcal{V}(\boldsymbol{x}_I,\boldsymbol{x}_{II})+\mathcal{U}_{II}(\boldsymbol{x}_{II},\boldsymbol{p}_{II})]}\mathrm{d}\boldsymbol{x}_{II}\mathrm{d}\boldsymbol{p}_{II}}{\int \mathrm{e}^{-\beta\mathcal{U}_{II}(\boldsymbol{x}_{II},\boldsymbol{p}_{II})}\mathrm{d}\boldsymbol{x}_{II}\mathrm{d}\boldsymbol{p}_{II}}\right)\end{aligned} \tag{3}$$

which is known as the "potential of mean force" or simply "mean field" (MF), $W(\boldsymbol{x}_I) \equiv \Delta A(\boldsymbol{x}_I)$. This is the field experienced by the explicitly treated molecules in a given configuration $\{\boldsymbol{x}_I\}$ due to the average interactions with the environment. Furthermore, we can assume that the non-bonded interactions can be separated into (long-range) electrostatic and (short-range) dispersion-repulsion contributions. Accordingly, we have for the interaction potential, $\mathcal{V}$:

$$\mathcal{V} = \mathcal{V}_{\mathrm{disp-rep}} + \mathcal{V}_{\mathrm{elec}} + \mathcal{V}_{\mathrm{cav}} \tag{4}$$

where the last term on the r.h.s. is a "hard wall" potential that has been introduced to account explicitly for the physical confinement of molecules belonging to system I inside the cavity of the macroscopic system. If we substitute Eq. 4 into Eq. 3, we can define, after simple mathematical manipulation, three corresponding terms for $W(\boldsymbol{x}_I)$:

$$W(\boldsymbol{x}_I) = W_{\mathrm{disp-rep}}(\boldsymbol{x}_I) + W_{\mathrm{elec}}(\boldsymbol{x}_I) + W_{\mathrm{cav}}(\boldsymbol{x}_I) \tag{5}$$

where

$$W_{\mathrm{cav}}(\boldsymbol{x}_I) = -kT \ln\left(\frac{\int \mathrm{e}^{-\beta[\mathcal{V}_{\mathrm{cav}}+\mathcal{U}_{II}]}\mathrm{d}\boldsymbol{x}_{II}\mathrm{d}\boldsymbol{p}_{II}}{\int \mathrm{e}^{-\beta\mathcal{U}_{II}}\mathrm{d}\boldsymbol{x}_{II}\mathrm{d}\boldsymbol{p}_{II}}\right) \tag{6}$$

$$W_{\mathrm{elec}}(\boldsymbol{x}_I) = -kT \ln\left(\frac{\int \mathrm{e}^{-\beta[\mathcal{V}_{\mathrm{elec}}+\mathcal{V}_{\mathrm{cav}}+\mathcal{U}_{II}]}\mathrm{d}\boldsymbol{x}_{II}\mathrm{d}\boldsymbol{p}_{II}}{\int \mathrm{e}^{-\beta[\mathcal{U}_{II}+\mathcal{V}_{\mathrm{cav}}]}\mathrm{d}\boldsymbol{x}_{II}\mathrm{d}\boldsymbol{p}_{II}}\right) \tag{7}$$

$$W_{\mathrm{disp-rep}}(\boldsymbol{x}_I) = -kT \ln\left(\frac{\int \mathrm{e}^{-\beta[\mathcal{V}_{\mathrm{disp-rep}}+\mathcal{V}_{\mathrm{elec}}+\mathcal{V}_{\mathrm{cav}}+\mathcal{U}_{II}]}\mathrm{d}\boldsymbol{x}_{II}\mathrm{d}\boldsymbol{p}_{II}}{\int \mathrm{e}^{-\beta[\mathcal{V}_{\mathrm{elec}}+\mathcal{V}_{\mathrm{cav}}+\mathcal{U}_{II}]}\mathrm{d}\boldsymbol{x}_{II}\mathrm{d}\boldsymbol{p}_{II}}\right) \tag{8}$$

In particular, W_{cav} is known as the cavitation free energy, i.e. the work spent to form a cavity into the liquid, which depends only on the cavity size and shape and is a constant for a fixed cavity. It should be noted that, rigorously speaking, a separation of the interaction energy into different contributions, as shown in Eq. 4, is not physically possible following a general quantum mechanical treatment. However, according to the Ben–Naim's definition of the solvation process [14], we can conveniently assume that the mean field, $W(\boldsymbol{x}_I)$, is composed of the conceptually simple terms provided in Eq. 5 (see, for example, the comprehensive review of Tomasi and Persico [1]).

Various continuum models have been proposed in the literature that differ in the way W is approximated [1, 15,16]. The present MF model is the result of a long-time effort to develop a sophisticated tool that allows to study efficiently solvent effects on generic solute molecules, with or without the presence of explicit solvent molecules and using different levels of theory, from computationally inexpensive, but less accurate, molecular mechanics (MM) methods to more realistic hybrid QM/MM or full QM methods. Moreover, such model allows to perform molecular dynamics both in Born–Oppenheimer [7,8] and extended-Lagrangian [9–12] frameworks. In the following sections, we describe in some detail how $W_{\mathrm{disp-rep}}$, W_{elec} and W_{cav} can be effectively modeled. Here, we want to sketch the MF formalism in the case that system I has a quantum core, e.g. the solute, which can be polarized in response to solvent effects. In analogy to the classical expression of the free energy, $A(\boldsymbol{x}_I) = \mathcal{U}_I(\boldsymbol{x}_I) + W(\boldsymbol{x}_I)$, we have for a quantum mechanical system:

$$A(\boldsymbol{x}_I) = E(\boldsymbol{P}, \boldsymbol{x}_I) + W(\boldsymbol{P}, \boldsymbol{x}_I) \tag{9}$$

where $E(\boldsymbol{P}, \boldsymbol{x}_I)$ and $W(\boldsymbol{P}, \boldsymbol{x}_I)$ are respectively the internal energy and the mean field contributions and $\boldsymbol{P}$ is the electronic density matrix. The explicit dependence of the latter two terms on $\boldsymbol{P}$ is to remark that they are mutually polarized and, as a consequence, the quantum mechanical solution should be obtained self-consistently. Moreover, when the explicit molecular system is treated by

a mixed quantum/classical potential (QM/MM), e.g. a solute (QM) embedded with MM solvent molecules, $E(\boldsymbol{P},\boldsymbol{x}_I)$ can be conveniently partitioned as proposed recently in [6]:

$$E(\boldsymbol{P},\boldsymbol{x}_I) = E^{\text{model,QM}}(\boldsymbol{P},\boldsymbol{x}_I)+E^{\text{real,MM}}(\boldsymbol{x}_I) - E^{\text{model,MM}}(\boldsymbol{x}_I) \quad (10)$$

where we refer to the entire molecular cluster as *real*, whereas *model* is the quantum core (solute) plus the point charges located at the same positions of the remaining explicit MM atoms (e.g. the set of the partial atomic charges of a non-polarizable MM force field). In Eq. 10, $E^{\text{model,QM}}(\boldsymbol{P},\boldsymbol{x}_I)$ is the internal energy of *model* treated at QM level, expressed as a function of the system coordinates, $\boldsymbol{x}_I$, and density matrix, $\boldsymbol{P}$, $E^{\text{real,MM}}(\boldsymbol{x}_I)$ is the MM internal energy of the *real* system, which obviously depends only on $\boldsymbol{x}_I$, and $E^{\text{model,MM}}(\boldsymbol{x}_I)$ is the MM internal energy of *model*. Although such decomposition is not derived straightforwardly, it provides a well-defined, single valued and differentiable potential well suited to perform QM/MM molecular dynamics simulations.

2.2 The cavitation free energy, W_{cav}

Given an explicit molecular system (system I), the first step of the mean field approach consists of creating a suitable cavity into the dielectric continuum. To this end, we have used an improved GEPOL procedure [17, 18] that allows to form efficiently cavities of general shape from interlocked spheres, typically centered on the atomic sites, and to partition the obtained cavity surface in small triangular elements (*tesserae*), which are needed to solve numerically the electrostatic problem (see Sect. 2.3). In particular, if the whole solvent is treated implicitly, the cavity surface is built up on the basis of the solvent accessible surface of the solute (usually in full QM calculations). On the other hand, if some solvent molecules, say the first two or three solvation shells, are treated explicitly along with the solute, then the cavity surface is more conveniently chosen as a smooth and regular surface, e.g. a sphere or an ellipsoid.

The cavitation free energy, W_{cav}, is defined as the work needed to form such a cavity into the solvent, regardless of the chemical nature of the "solute" (system I). Hence, for a given solvent under specific physical conditions, e.g. density and temperature, it does depend only on the size and shape of the cavity and is a constant for a fixed cavity. From a physical point of view, the actual value of W_{cav} is the result of the subtle balance of enthalpic and entropic contributions due to solvent–solvent interactions that, macroscopically, give rise to the surface tension. Although several different models have been proposed in the literature to evaluate the cavitation free energy, in the context of continuum methods the most common are the hard sphere model, as applied in Pierotti's scaled particle theory (SPT) [19] and its generalization to cavities formed by a collection of spheres, known as the Claverie–Pierotti (CP) model [20]. Recently, a general and quite flexible approach [21] to compute W_{cav}, based on the SPT of a liquid of spherocylindric molecules, has been presented as an extension of the usual formulation. Such a model is particularly well-suited in the more general case of non-spherical (rod-like) solvents and improves the description of the solvation process when solvent is not isotropically distributed, for example in liquid crystals where solvent can be highly ordered [21].

In the present work, all the solvents and cavities considered are nearly spherical and, for convenience, we have adopted the popular CP model [20], as implemented in Gaussian03 [22]. Also, when part of the solvent was treated explicitly, an "hard wall" potential has been used to keep the molecules confined into a spherical cavity (solvent molecules approaching the cavity boundaries undergo elastic collisions with respect to their center of mass motion).

2.3 The electrostatic (reaction field) free energy, W_{elec}

According to electrostatic theory [23], a charge distribution embedded in the cavity of a dielectric continuum experiences an electric field, commonly known as "reaction field", due to the induced polarization of the dielectric medium. Within the continuum approximations, the reaction field potential, Φ_{RF}, can be obtained by solving the Poisson equation [23] both analytically, if the cavity shape is a sphere or an ellipsoid, or numerically with either the finite difference [24,25] or the boundary element (BE) [26–29] methods. The so-called polarizable continuum model (PCM) is probably the most refined and reliable of all the BE approaches both for quantum mechanics (QM) and molecular mechanics (MM) applications, in a wide range of problems ranging from structure to thermodynamics, kinetics and spectroscopy in both isotropic and anisotropic environments [1,3,30]. The method has been recently improved to take into proper account also the so called outlying charge effects (due to the solute electronic density tails that penetrate into the continuum) [3,31,32,37]. A quite popular variant of the method is the conductor-like model [38,39] (corresponding to the PCM limit for infinite dielectric constant, but used with satisfactory results also for solvent of rather low polarity): its implementation in the PCM framework (CPCM) is described in [39,40]. In all

these methods the solute or the region of interest is represented in atomic details and embedded in a suitable cavity, whose surface is finely subdivided into small tiles (*tesserae*). Next, the solvent reaction field is described in terms of apparent point charges centered on tesserae ($q_{\rm asc}$) and self-consistently adjusted to the solute electron density [41].
The solute–solvent interaction energy is

$$E_{\rm int} = \Phi^{\dagger}\mathbf{q_{asc}} = \sum_{i}^{\rm tesserae} \Phi_i q_{{\rm asc},i} \tag{11}$$

where the column vectors Φ and $\mathbf{q_{asc}}$ collect the solute electrostatic potential and the solvation charges in the surface tesserae, respectively. The solvation charges are computed as $\mathbf{q_{asc}} = \mathbf{Q}\Phi$, where $\mathbf{Q}$ is a geometric matrix, depending on the position and size of the surface tesserae. The core of the model is then the definition of the $\mathbf{Q}$ matrix. For isotropic solvents it is defined as [41]

$$\mathbf{Q} = \mathbf{T^{-1}R}$$

$$\mathbf{T} = \frac{\epsilon+1}{\epsilon-1}\mathbf{S} - \frac{1}{2\pi}\mathbf{DAS} \tag{12}$$

$$\mathbf{R} = -\mathbf{I} + \frac{1}{2\pi}\mathbf{DA} \tag{13}$$

where ϵ is the dielectric constant, $\mathbf{I}$ is the unit matrix, $\mathbf{A}$ is a diagonal matrix collecting the area of tesserae and the matrices $\mathbf{S}$ and $\mathbf{D}$ are related to the electrostatic potential and to the electric field generated by the solvation charges, respectively. Taking the limit for $\epsilon \to \infty$, the CPCM, much simpler, expression is obtained:

$$\mathbf{Q}_{\infty} = -\mathbf{S}^{-1} \tag{14}$$

$$\mathbf{q}_{\infty} = \mathbf{Q}_{\infty}\Phi \tag{15}$$

where

$$S_{ii} = 1.0694\sqrt{\frac{4\pi}{a_i}} \tag{16}$$

$$S_{ij} = \frac{1}{|\boldsymbol{s_i} - \boldsymbol{s_j}|} \tag{17}$$

$\boldsymbol{s_i}$ and a_i being the position vector and the area of the ith tessera, respectively. To model a real solvent, with finite dielectric constant, the solvation charges have to be scaled:

$$\mathbf{q_{asc}} = \frac{\epsilon-1}{\epsilon}\,\mathbf{q}_{\infty} \tag{18}$$

Hence, for a given molecular configuration of the explicit system, $\boldsymbol{x}_I$, the electrostatic free energy, $W_{\rm elec}$, are given by

$$W_{\rm elec} = \frac{1}{2}\sum_{i} \Phi_i(\boldsymbol{r_i}) q_{{\rm asc},i} \tag{19}$$

We recall that the PCM formulation originally derived by Tomasi and coworkers [26] (in the PCM literature this version is usually referred to as DPCM) is less accurate than those illustrated above, because it does not include the effects of the outlying charge. However, when solutes are treated at the MM level electronic tails are obviously not present, and PCM reduces to DPCM. Note that the DPCM solvation charges depend on the normal component of the electric field on the surface (unlike PCM and CPCM charges depending on the potential)

$$\left[2\pi\frac{\epsilon+1}{\epsilon-1}\mathbf{A}^{-1} - \mathbf{D}^{\dagger}\right]\mathbf{q_{asc}} = -\mathbf{E_n} \tag{20}$$

where $\mathbf{D}^{\dagger}$ is the adjoint of matrix $\mathbf{D}$ and $\mathbf{E}_n$ collects the normal electric field on tesserae. This implies an increased computational burden and some numerical noise, unless very small tesserae are used. Thus the CPCM version, together increasing effectiveness and numerical stability, allows to use the same model in MM, QM and QM–MM applications without introducing any inconsistency. Indeed, we have found that $W_{\rm elec}$ converges more rapidly with the number of tesserae, $N_{\rm tes}$, and generally is more stable by using the present methodology (data not shown). On the grounds of the same tests, we have chosen $N_{\rm tes}$ in such a way that the average tessera area was less than 0.7 Å^2 for all the systems discussed in the following where part of the solvent was treated explicitly (spherical cavities), and about 0.2 Å^2 when the cavity was constructed on the solute molecular surface (all solvent implicitly treated). Also, in the former case the dielectric continuum response, e.g. the reaction field, was calculated considering a radius of $R_{rf} = R_{\rm cav} + 1.8$ Å, with $R_{\rm cav}$ the radius of the spherical cavity, to avoid numerical instabilities.

The cavity surface has been partitioned using the GEPOL procedure [17,18] and the $\mathbf{S^{-1}}$ matrix has been obtained by direct inversion and stored once for ever at the beginning of each simulation since in the present approach $\mathbf{S}$ does not depend on the dynamics of the explicitly simulated system. For the same reason, the energy derivative with respect to a generic coordinate α is much simpler than the complete expressions including cavity deformation [39], reducing to

$$W^{\alpha}_{\rm elec} = \sum_{i}^{\rm tesserae} \Phi^{\alpha}_i q_{{\rm asc},i} \tag{21}$$

Of course, computation of the derivative of the electrostatic potential generated by the explicitly simulated system at tessera i (Φ^{α}_i) is straightforward whenever one

deals with classical systems described in terms of fixed atomic charges.

2.4 The dispersion–repulsion free energy, $W_{\mathrm{disp-resp}}$

The dispersion–repulsion contribution, $W_{\mathrm{disp-resp}}$, is related to short-range interactions between system I and the outer solvent. Generally, $W_{\mathrm{disp-resp}}$ is determined from some empirically derived parameters and, in the case of a quantum mechanical treatment of the explicit system, does not enter in the self-consistent calculations of the wavefunction, $\Psi(\boldsymbol{x}_I)$, which is only perturbed by the long-range electrostatic effects. According to usual methods [42,43] developed in the PCM framework, we can assume that the dispersion–repulsion interactions can be modeled by a simple empirical potential and, consequently, we can obtain $W_{\mathrm{disp-resp}}$ by integrating such potential times the density distribution of the outer solvent, treated as a continuum, from the cavity surface to infinity. This is a very convenient and effective procedure to determine the dispersion–repulsion free energy when the whole solvent is treated implicitly. In this case, $W_{\mathrm{disp-resp}}$ represents effectively the average solute–solvent(implicit) interactions. However, if the solute and a few solvent shells are treated explicitly with the cavity surface chosen as a smooth and regular surface, then the dispersion–repulsion term accounts mainly for the short-range solvent(explicit)–solvent(implicit) interactions and is more conveniently modeled by the effective approach originally presented in [4] (see also [5,6] for applications in the context of MM and QM/MM molecular dynamics simulations, respectively) and described in the following.

First, we assume that the solute center of mass position is constrained at the center of the cavity, whereas solvent molecules are confined by means of an "hard wall" potential (see Sect. 2.2). Note that for consistency with [4], hereafter we refer to the dispersion–repulsion term as the van der Waals (VdW) potential, $W_{\mathrm{vdw}} \equiv W_{\mathrm{disp-resp}}$, formally obtained following the same statistical mechanical derivation presented above (Eq. 8). W_{vdw} should be regarded as a short-range potential, i.e. acting solely on the molecules very close to the cavity boundaries. Indeed, we can express W_{vdw} in terms of a potential that depends on the distance of a molecule from the cavity surface, defined as the vector pointing from specific site(s) on the molecule to its closest point on the surface. Also, we can reasonably assume that such a potential does not change significantly for cavities of sufficiently regular shapes containing the same molecular systems at the same physical conditions, in other words we can assume that it does not depend on the local curvature of the surface. Then, we can set up an empirical procedure to determine the functional form of such a potential by considering, for convenience, a spherical cavity.

First, we assume that the potential $W_{vdw}(r)$ is a radial function acting only on the center of mass of the explicit solvent molecules embedded in the spherical cavity. Then, we express $W_{vdw}(r)$ as the sum of a set of N_g Gaussian functions, g_i, whose centers, r_i, are equally spaced over a radial direction of the spherical cavity:

$$W_{vdw}(r) = \sum_{i}^{N_g} g_i(r - r_i) = \sum_{i}^{N_g} \lambda_i e^{\left[-\frac{(r-r_i)^2}{2\sigma^2}\right]} \tag{22}$$

Also, we assume that each $g_i(x)$ function has the same spread, σ, but a variable height, λ_i, which is a multiple of a fixed amount ($\lambda_i = n_i h$) and is determined empirically from a test simulation using the overall density, ρ_0, of the system as a reference. The method works as follows. Given a spherical cavity with radius R_s, the number of Gaussian functions is obtained from $N_g = R_s/(2\sigma)$ (with $\sigma = 0.125$ Å), that is, $g_{i+1}(x)$ is centered at a distance 2σ from $g_i(x)$ along the radial direction. Moreover, we divide ideally the sample into N_g concentric spherical layers, each one corresponding to a $g_i(x)$ function, with a width of 2σ. During the test simulation, after a certain time interval, τ (with $\tau = 20$ ps), the local densities $\rho(r_i)$ computed on each spherical layer are compared to the overall density ρ_0, which is a constant. If $\rho(r_i) > \rho_0$, the corresponding Gaussian height centered at r_i is increased by one unit h, with $h = 0.01$ kJ/mol, ($n_{i,\mathrm{new}} = n_{i,\mathrm{old}} + 1$), otherwise, if $\rho(r_i) < \rho_0$, the height is decreased by the same amount ($n_{i,\mathrm{new}} = n_{i,\mathrm{old}} - 1$). In a few nanoseconds, $W_{vdw}(r)$ converges to an optimal potential function, that is, it does not change significantly anymore. The final form of the van der Waals term is determined as an average of the instantaneous $W_{vdw}(t_i)$ collected during such a test simulation.

It should be noted that $W_{vdw}(r)$, obtained as described above, depends on the specific physical conditions, such as density and temperature, and on the solvent molecular model considered and is not generally transferable to other solvents. However, we have observed that for water, the popular SPC and TIP3P models give rise to rather similar potentials (data not shown).

2.5 Calculation of spectroscopic properties

Spectroscopic parameters of molecules in solution represent invaluable sources of information on both their properties and their interactions with the environment (solvent). In most cases, the influence of the solvent on spectroscopic properties could not be neglected. To conveniently describe such solvent effects, we can

conceptually make a decomposition in direct and indirect contributions: the former are due to "direct" solute–solvent interactions, such as electrostatic or hydrogen bond interactions, whereas the latter are responsible for changes in the average structure and in the dynamics of the solute, which ultimately could affect solute spectroscopic parameters. Both contributions should be properly taken into account by theoretical models developed to reproduce accurately spectroscopic observables. In particular, there are a number of spectroscopic parameters, e.g. NMR shielding constants, [44], UV-vis electronic transitions [45] and EPR hyperfine coupling constants [46,47], that do depend on the solute–solvent structural configuration, or better on the statistical average of such configurations, and other parameters that are also related to the detailed dynamical behaviour of the molecular system, such as infrared spectra and, more generally, spectroscopic parameters related with molecular relaxation.

A very effective theoretical approach to study all such kind of spectroscopic properties is to use molecular dynamics methodologies combined with high level quantum mechanical calculations. As far as solute–solvent systems are considered, we have shown [48,49] that a computational protocol both flexible and reliable consists of (1) sampling a representative number of (solute–solvent) molecular configurations by means of MD simulations using the mean field model described above and (2) performing a posteriori spectroscopic calculations on the same explicitly simulated system, including, for consistency, solvent effects provided by the dielectric continuum. About such a procedure, a few points deserve some comments. In the first step, we should carefully consider if the solvent can be fully treated implicitly or if some solvent shells have to be included explicitly in the simulation. Generally speaking, for polar and protic solvents, like water, capable of forming specific solute–solvent interactions, it is necessary to treat explicitly at least part of the solvent. On the other hand, for weakly interacting non-polar solvents, an implicit description is, in most cases, efficient and accurate at the same time. Moreover, whenever the solvent is treated explicitly, it is usually more convenient to model it at a lower level of theory with respect to the solute, e.g. by using hybrid QM/MM approaches. Note, however, that although such a choice is very advantageous for the sampling, it could be inadequate for the subsequent quantum mechanical calculations of spectroscopic properties. It has been shown that in some cases, e.g. NMR chemical shifts, subtle quantum effects, such as quantum exchange and Pauli exclusion principle, between the solute and the solvent molecules in the first solvation shell are relatively important and should be considered in the calculations (see, for example [50]). Hence, if necessary, a limited number of solvent molecules, along with the solute, are treated at QM level, whereas the remaining part of the solvent is included in the effective quantum Hamiltonian as an embedding electrostatic field, made up of point charges. In addition, the dielectric continuum is also included to account for long-range interactions. Furthermore, the QM computational method used in step (2) (spectroscopic calculations), including the choice of the basis set, is completely independent and could differ from the one used for the sampling. Remarkably, depending on the parameter considered, the present protocol allows to select the most appropriate QM methodology [51]. On the other hand, state of the art methodologies based on direct "on the fly" calculations of spectroscopic parameters, such as UV spectra, [52] from ab initio MD simulations within the CPMD [53] formalism are generally restricted to less satisfactory DFT methods relying on the generalized gradient approximation (GGA).

2.6 Computational details

All the room temperature QM and QM/MM molecular dynamics simulations reported in the present work were performed according to the atom centered density matrix propagation (ADMP) [10–12] formalism, in which the density matrix of the atomic basis set evolves together with the nuclei as dynamic variable. The core and valence orbitals were weighted differently during the dynamics with $\mu_{\text{valence}} = 0.1\,\text{amu}\,\text{bohr}^2 \approx 180\,\text{a.u.}$ for the valence electrons and μ_{core} obtained according to the tensorial fictitious mass scheme described in [11]. A constant thermal energy has been enforced by scaling nuclear velocities every 2,500 steps, with a time step of 0.2 fs. The DFT B3LYP [54] method based on Becke's hybrid exchange–correlation functional [13] was used in all simulations, as well as quantum mechanical calculations, with an ad hoc atomic basis set, hereafter referred to as N06, that reduces significantly the basis set superposition error (BSSE) and is particularly well-suited for ab initio molecular dynamics, with the exception of formamide for which a B3LYP/6-31G(d,p) level of theory was employed. Simulations in the condensed-phase were performed using the MF model described in the preceding sections. In the aqueous solutions simulations of acetone, acrolein and glycine radical, the solvent was partially included explicitly according to an hybrid QM/MM scheme: the QM region, which included only the solute computed at B3LYP level of theory, was solvated with about 130 TIP3P water molecules (MM region) and embedded into a spherical cavity of a dielectric medium with a radius of 11.8 Å. The center of mass

of the solute was constrained at the center of the cavity. Interactions between QM and MM parts, as usual, were composed of both electrostatic and van der Waals interactions, where the latter were modeled according to the Lennard–Jones potential parameterized by Freindorf et al. [55] [note that therein a different basis set was used, 6-31+G(d), which has provided very similar results in test calculations compared to the basis set (N06) employed in this work]. In the simulation of formamide in acetonitrile, solvent was included implicitly and the molecular cavity was obtained according to the GEPOL procedure using Pauling atomic radii. All the quantum mechanical calculations and the QM and QM/MM molecular dynamics simulations have been performed with a modified version of the Gaussian package [22].

3 Results and discussion

3.1 NMR shielding constants

Nuclear magnetic resonance (NMR) spectroscopy is probably the most popular and powerful spectroscopic technique in chemistry. Detailed information on molecular structure and dynamics could be extracted from NMR mono and bidimensional spectra. As an example, NMR chemical shifts of magnetically active isotopes, such as ^{1}H, ^{13}C or ^{17}O, are routinely used to characterize chemical compounds [56] and NMR relaxation measurements provide unique dynamical information of internal motions in proteins [57]. From a theoretical point of view, several studies [56,58] have shown that NMR shielding tensors can be computed quite accurately according to the gauge-including atomic orbital (GIAO) [59,60] formalism at both DFT or post-Hartree Fock level of theory. However, given the time resolution of NMR spectroscopy, dynamical averaging of molecular rotations and vibrations should often be taken into account in order to reproduce quantitatively experimental measurements.

Here, we consider, as a test case, an aqueous solution of acetone. If we evaluate the solvent effects on the ^{13}C and ^{17}O NMR isotropic shielding constants of its characteristic carbonyl group from quantum mechanical calculations of acetone and acetone–$(H_2O)_2$ cluster plus PCM, we obtain significant deviations from experimental results [$\Delta\sigma(^{17}O)$: 92.3 ppm (theory), 75.5 (exp.); $\Delta\sigma(^{13}C)$: −17.2 ppm (theory), −18.9 (exp.)]. Such discrepancies are largely due to the neglect of thermal fluctuations of the solute–solvent arrangement and improvements of the computed parameters can only be expected by inclusion of such dynamical information. In [49], we have shown how the combination of the MF model with high-level spectroscopic calculations leads to results with satisfactory accuracy for such a molecular system. In particular, the MF model has been used to simulate acetone with some explicit water molecules and, from the obtained configurational sampling, a posteriori quantum calculations of NMR shielding constants have been performed. Note that water was not entirely modeled as a continuum to account for the specific hydrogen bond interactions with the carbonyl group of acetone. In the present review, we compare the results obtained from the 1 ns classical MD simulations of [49], where the solute and the solute–solvent interactions have been parametrized on the basis of ab initio calculations, a short CPMD [53] simulation and experimental data, with a more sophisticated QM/MM simulation (see Sect. 2.6 for details).

Table 1 Average geometrical parameters of acetone in aqueous solution issuing from MM and QM/MM simulations

	MM[a]	QM/MM
C=O	1.222 (0.002)	1.226 (0.003)
C–C	1.511 (0.002)	1.508 (0.003)
C–H	1.090 (0.001)	1.098 (0.003)
C–C=O	121.5 (0.5)	121.0 (0.5)
C–C–C	115.7 (0.6)	117.7 (0.5)

Bond distances are in Å, angles in degrees. Standard error are reported in parentheses
[a] Taken from [49]

First, we analyze the most relevant aspects of the solute–solvent structure. In Table 1, the average geometrical parameters of acetone in aqueous solution are reported for both MM and QM/MM simulations. An overall satisfactory agreement is observed. Note that the extent of the C=O bond distance was readjusted in MM by choosing a different equilibrium distance with respect to the gas phase ($r_0 = 1.2080$ Å, gas phase; $r_0 = 1.2170$ Å, solution), whereas the carbonyl elongation observed in QM/MM comes out naturally from the interactions with water [$d(C = O) = 1.213$ Å, gas phase optimized structure]. Also, very close results are obtained considering the radial distribution functions (RDF) of the water molecules with respect to the carbonyl oxygen, as shown in Fig. 2. From the integration of the RDF's, it has been found that, as expected, two water molecules are on average hydrogen bonded to the carbonyl group, where each water molecule forms only one hydrogen bond with the acetone oxygen, while the second hydrogen atom is pointing away (see Fig. 2b). Moreover, a very similar hydrogen bond distribution has been obtained in the two simulations, as reported in Table 2. All the results suggest that the two MD methodologies considered (MM and QM/MM) are nicely consis-

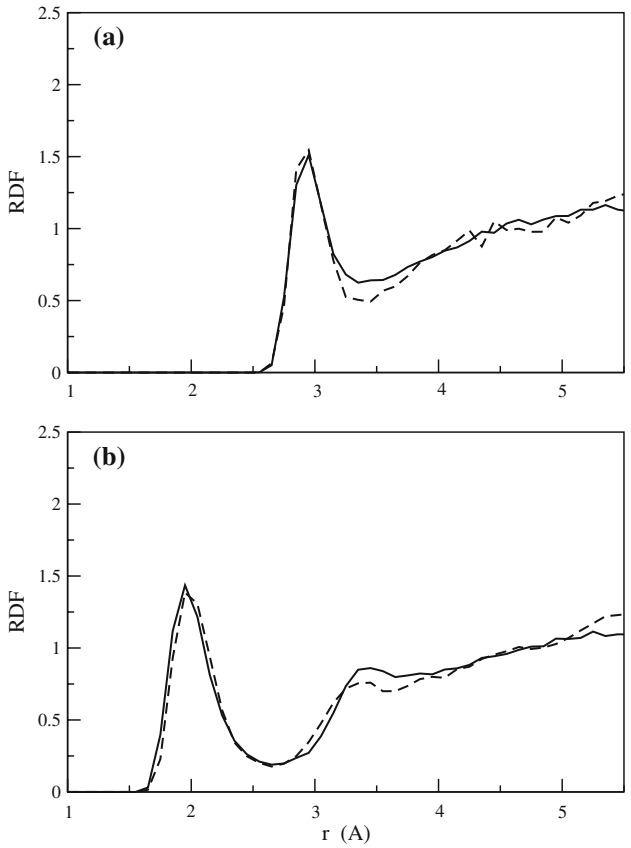

Fig. 2 **a** O· · ·Ow and **b** O· · ·Hw radial distribution functions of acetone in aqueous solution as obtained from MM [49] (*solid line*) and QM/MM (*dashed line*) molecular dynamics simulations

Table 2 Average number of hydrogen bonds between acetone and water issuing from MM and QM/MM simulations

	MM[a]	QM/MM
No. of H-bonds	2.0	2.0
% 1 H-bond	15	13
% 2 H-bonds	66	74
% 3 H-bonds	19	13

[a] Taken from [49]

tent to one another. Note, however, that MM generally requires a careful development and validation of a reliable solute–solvent force field, introducing, if necessary, additional non-standard parameters, such as the lone pairs on the carbonyl oxygen (neglecting such terms could result in an erroneous solute–solvent structural arrangement, see for example [49]).

In Table 3, the NMR isotropic shielding constants relative to the acetone C=O group, both in gas phase and in aqueous solution, are reported and compared to available experimental data. QM/MM and MM [49] results were obtained according to the computational

Table 3 ^{13}C and ^{17}O NMR isotropic shielding constants of acetone in the gas phase and in aqueous solution, computed within the gauge-including atomic orbital (GIAO) formalism at B3LYP/6-311+G(2d,2p) level of theory

		MM[a]	QM/MM	Exp.
^{17}O	Gas phase	−340	−342.8	
	Solution	−253	−254.6	
	Δ	87	88.2	75.5
^{13}C	Gas phase	−23	−29.6	
	Solution	−42	−48.3	
	Δ	−19	−18.7	−18.9

Values are in ppm, standard error is 2–5 ppm
[a] Taken from [49]

procedure described in Sect. 2.5, including in the QM region the two closest water molecules to the carbonyl group. Note that QM/MM gas phase shielding constants were computed from the acetone optimized structure. Remarkably, the extent of solvent effects on such magnetic properties is reasonably well reproduced in both cases. The changes in the ^{13}C NMR shieldings on-going from the gas phase to the liquid are in excellent agreement with experiment, within the statistical noise. On the other hand, results on ^{17}O NMR shieldings are somewhat less accurate. However, an error of about 10 ppm should be considered acceptable and, besides, no scaling factors or other kind of corrections were applied to the reported results.

3.2 UV electronic transitions

Most carbonyl compounds show a characteristic UV absorption corresponding to the excitation of a non-bonding (n) electron on oxygen into an antibonding $\pi*$ orbital of the C=O group in the region between 4.0 and 4.5 eV. Also, a solvatochromic blueshift of such a transition is usually observed in going from the gas phase to aqueous solution, as found for acetone [61–63] and predicted for formaldehyde. [64,65] The generally accepted interpretation of such a shift is that the electronic ground-state has a larger dipole moment with respect to the first excited state and, as a consequence, it is more stabilized in polar solvents, such as water. However, the actual extent of the observed blueshift is the result of different and opposite effects, not only polar effects (see for example [66]).

In ref. [48] the computational approach described in the Sect. 2 was fruitfully employed to understand more deeply such physical phenomenon by considering in some detail the case of acrolein (CH_2CHCHO), which is the simplest representative of the α,β-unsatured aldehyde class. In particular, a NVT QM/MM simula-

tion of acrolein + 134 TIP3P water molecules and a full QM simulation of acrolein in vacuo were performed for 24 ps, including 4 ps of equilibration, using the ADMP methodology [10–12] (see ref. [48] for details). UV excitation energies and oscillator strengths were computed within the TD-DFT formalism employing the B3LYP functional and the 6-311++G(2d,2p) basis set. The consistency of such basis set in spectroscopic calculations was validated in a recent work [67].

The nature of the solvent (water) effects on the UV $n \rightarrow \pi*$ transition energy of acrolein was analyzed by evaluating the relative contributions due to direct (solvent polarization and H-bonding) and indirect (solute structural rearrangements) effects. Note that the UV absorption spectrum of acrolein was studied experimentally in different solvents [68–74], as well as in gas phase [75–77], and a solvatochromic blueshift of 0.20–0.25 eV was observed, as a result of water solvation. In Table 4, results for the computed blueshift of the $n \rightarrow \pi^*$ vertical transition of acrolein, from the gas phase and condensed phase MD simulations at room temperature, are reported and in Fig. 3 the corresponding spectra are plotted. The overall computed blueshift is 0.26 ± 0.01 eV (last line of Table 5), in good agreement with experiments (0.20–0.25 eV). Such a result was obtained by treating, in the aqueous solution calculations, the two closest water molecules to the C=O group at QM level, along with acrolein (Acrolein + $2H_2O^{QM}$ + $132H_2O^{MM}$ + MF). Besides, when water molecules are all treated as point charges (Acrolein + $134H_2O^{MM}$ + MF), the blueshift is unchanged within the statistical noise, 0.25 ± 0.01 eV, as noted in a recent study for the case of acetone [49]. This means that solvent effects on the $n \rightarrow \pi*$ vertical transition are essentially of electrostatic nature. Also, we have evaluated the separate contributions to the blueshift coming solely from the solute structural changes (second line of Table 5) and from the first solvation shell of the C=O group (Acrolein + $2H_2O^{QM}$), that is including the closest two water molecules to the carbonyl oxygen. Remarkably, solute geometry distortions lead to a not negligible redshift (−0.08 eV) and, hence, the direct solvent effects on the spectroscopic property, once the solute geometry has changed, is about 0.34 eV. Also, about half of such amount is provided by the first two water molecules surrounding the C=O group (0.18 eV), which means that H-bonding and bulk effects are nearly equal.

3.3 EPR isotropic coupling constants

It is now well established that, contrary to the parent molecule, glycine radical prefers a neutral structure (i.e. NH_2-CH-COOH in place of NH_3^+-CH-COO^-) even in aqueous solution [78,79] and that cationic and anionic forms play a negligible role except at extreme pH values. However, some of the isotropic hyperfine splittings of the radical (especially H^α) show values quite far from those expected for similar radicals or observed for the zwitterionic form of the radical in the solid state [80]. This stimulated a number of theoretical studies, which ended up with a satisfactory explanation of the EPR parameters in structural terms when intramo-

Table 4 Average geometrical parameters of acrolein issuing from gas phase (QM) and aqueous solution (QM/MM) simulations

	Gas phase	Solution	Δ
C=O	1.214 (0.002)	1.224 (0.002)	0.010
C–C	1.479 (0.002)	1.465 (0.002)	−0.014
C=C	1.339 (0.002)	1.341 (0.002)	0.003
C–C=O	123.9 (0.3)	124.0 (0.3)	–
C=C–C	121.0 (0.3)	120.6 (0.3)	–
C=C–C=O	180 (1)	180 (1)	–
μ	3.37 (0.01)	4.97 (0.01)	1.6

Bond distances are in Å, angles in degrees. Standard errors are reported in parentheses. Data taken from [48]

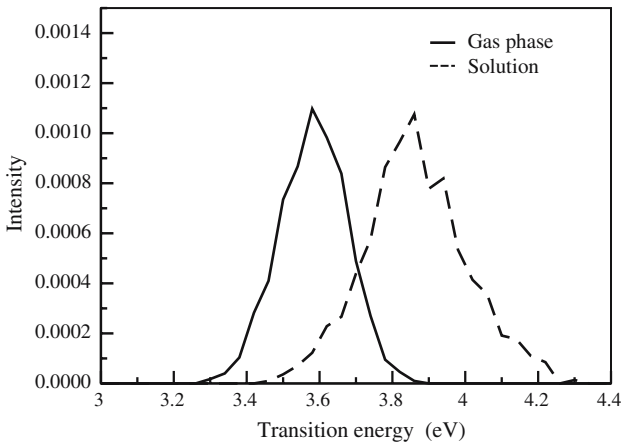

Fig. 3 Acrolein UV spectra relative to the first singlet $n \rightarrow \pi^*$ electronic transition, as computed from the gas phase and aqueous solution MD simulations at TD-DFT B3LYP/6-311+G(2d,2p) level of theory

Table 5 UV $n \rightarrow \pi^*$ transition energies of acrolein in the gas phase and in aqueous solution, computed within the TD-DFT formalism at the B3LYP/6-311+G(2d,2p) level of theory

	$n \rightarrow \pi^*$ (eV)
Gas phase	3.58
Solution	
Acrolein	3.49 (−0.08)
Acrolein + $2H_2O^{QM}$	3.68 (+0.10)
Acrolein + $2H_2O^{QM}$+$132H_2O^{MM}$+PCM	3.84 (+0.26)

Values are in eV. Standard error is 0.01 eV. Data taken from [48]

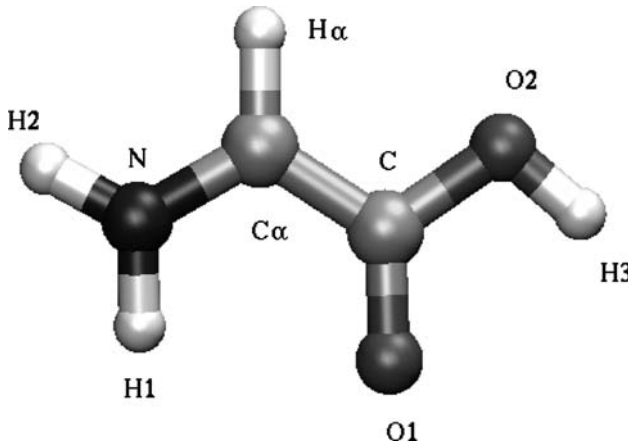

Fig. 4 Optimized structure of the glycine radical computed at B3LYP/NB level of theory

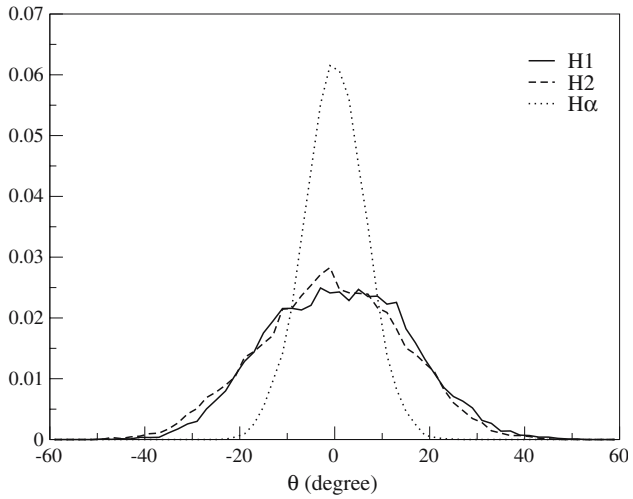

Fig. 5 Normalized probability distributions of the ϕ(H1NC$^{\alpha}$C), ϕ(H2NC$^{\alpha}$C) and ϕ(NH$^{\alpha}$CC$^{\alpha}$) dihedral angles of glycine radical in aqueous solution issuing from a QM/MM molecular dynamics simulation

lecular vibrational averaging and solvent effects were taken into account by an integrated computational procedure [81]. However, some limitations are still present in the theoretical approach since dynamical solvent effects were not taken into account and some hypothesis about additivity of different effects was unavoidable. The dynamical procedure described in this paper allows a more direct treatment in which all short-time dynamical effects are taken into account simultaneously in a coherent way. In analogy with the previous cases of acetone and acrolein, a NVT QM/MM simulation of the glycine radical solvated with 134 TIP3P water molecules was performed at room temperature (300 K) for 32 ps, including 4 ps of equilibration. The optimized gas phase structure of the glycine radical (see Fig. 4), computed at the B3LYP/NB level shows that the aminic group is to

Table 6 EPR hyperfine coupling constants calculated at B3LYP/EPR-II level of theory for the glycine radical (GlyR) in the gas phase, solvated with four water molecules (GlyR + 4H_2O) and by PCM, and in QM/MM aqueous solution simulation (see Fig. 4 for atom labels)

	GlyR	GlyR + 4H_2O	GlyR + 4H_2O + PCM	Simul.	Exp.
a(N)	5.45	4.07	4.19	5.58	6.38
a(H1)	−5.77	−8.95	−9.10	−5.70	−5.59
a(H2)	−3.73	−9.23	−9.39	−5.42	−5.59
a(C$^{\alpha}$)	11.85	6.70	6.22	10.90	
a(H$^{\alpha}$)	−14.54	−12.00	−11.74	−12.12	−11.77

some degree pyramidal, with ϕ(H1NC$^{\alpha}$C) = −9.5° and ϕ(H2NC$^{\alpha}$C) = −165.0°, whereas H$^{\alpha}$ is almost on the same plane of the N-C$^{\alpha}$-C group [ϕ(NH$^{\alpha}$CC$^{\alpha}$) = 1.6°]. On the other hand, in aqueous solution the glycine radical is, on average, approximately planar: the fluctuations of the above dihedral angles are symmetrically distributed around the planar conformation and the aminic hydrogens show an equal and broader distribution than H$^{\alpha}$ (see Fig. 5). Also, the first solvation shell has been analyzed in some detail. Considering H1, H2, H$^{\alpha}$ and H as hydrogen bond donors and N, O1 and O2 as hydrogen bond acceptors (see Fig. 4 for atom labels), we have obtained the following average number of hydrogen bonds formed during the aqueous solution simulation: H1 = 0.58, H2 = 0.68, N = 0.07, H$^{\alpha}$ = 0.23, O1 = 2.04, O2 = 0.87 and H = 1.00. Those values were obtained based on cut-offs for the donor–hydrogen–acceptor angle (60°, zero corresponding to the linear donor–hydrogen–acceptor arrangement) and the hydrogen–acceptor distance (2.6 Å). At the same time, full geometry optimization of a cluster containing glycine radical and four water molecules provides a structure very close to that shown in Fig. 4 of [81].

The isotropic hyperfine couplings computed by different models of the glycine radical are compared in Table 6 with the results averaged over the QM/MM trajectory. From the one hand, the remarkable agreement between the computed values issuing from our QM/MM simulation and experiment for all the available hyperfine splittings points out, once again, the reliability of the computational approach. From the other hand, the non-negligible difference between the results of the simulation and those obtained for the isolated glycine radical or the optimized cluster including the whole first solvation shell point out the role of solvent effects both from a static and a dynamic point of view. Starting from the quite disappointing results obtained for the isolated radical, inclusion of the first solvation shell leads to nearly equivalent H1 and H2 atoms, but the quantitative values remain quite far from experiment. Next, inclusion of

bulk solvent effects by the polarizable continuum model (PCM) [41] has a negligible effect, and only dynamical averaging (both intra- and inter-molecular) restores agreement with experiment. Thus, none of the static models is sufficient to provide even semi-quantitative results and a dynamical treatment is needed to provide a coherent picture. While the good results obtained in [81] suggest that in this case intra-solute dynamics plays a dominant role, the examples discussed in the previous sections show that this is not always the case, and point out the interest of an integrated approach including all the short-time dynamical effects (i.e. solute large amplitude vibrations and solvent librations).

3.4 Vibrational analysis and IR spectra

Formamide is the simplest molecule containing the N-C=O group found in peptides, and, for this reason, its spectroscopic properties have been extensively studied, both theoretically and experimentally. We investigated the potentialities of a dynamical analysis to characterize the vibrational motion of formamide in acetonitrile solution, when the solvent is implicitly included in the simulation. Solvent effects have been evaluated by comparison with a corresponding dynamics obtained for the system in the gas phase, and with a quantum vibrational analysis exploiting the same ab initio potential. Being the vibrational modes simulated at finite temperature, a reliable comparison with the quantum treatment requires for this latter a proper inclusion of anharmonicity. In this case we exploited a vibrational Hamiltonian in which cubic and quartic terms of the potential expansion have been included via a perturbative methodology [82,83].

Briefly, the dynamical analysis consisted in evaluating generalized normal modes Q_i and associated frequencies ν_i from eigenvectors and eigenvalues of the covariance matrix of the nuclear linear momenta. We recall that fundamental frequencies, as obtained by the quantum treatment, correspond to single excitations of vibrational quanta from the ground state, situated at the zero point energy of the molecule. On the other hand, the present dynamical analysis assigns characteristic frequencies of classic coupled anharmonic oscillators. As a consequence, frequency values are not constants of motion. On the contrary, the values depend upon the potential shape explored according to the configurational distribution, which, in turn, depends upon the total kinetic energy and, ultimately, on the temperature. Despite of the deep difference between the the two descriptions, the frequency values obtained by the two approaches reported in Table 7 can be nicely compared,

Table 7 Comparison of frequency values (cm^{-1}) for formamide in the gas phase obtained from the quantum mechanical pertubative approach and the first-principle MD trajectory analysis at B3LYP/6-31G(d,p) level of theory

Mode	Pert. method	MD
Out of plane NH_2	–	421
Wagging out of phase NCO/NH_2	564	561
Torsion	596	586
Out of plane CH	1,022	1,022
Wagging in phase NCO/NH_2	1,055	1,042
CN stretch	1,212	1,244
In plane COH sciss	1,405	1,389
In plane NH_2 sciss	1,582	1,605
CO stretch	1,775	1,758
CH stretch	2,814	2,822
Sym NH stretch	3,413	3,402
Asym NH stretch	3,525	3,555

Table 8 Frequency values and solvent shifts (cm^{-1}) for formamide in the gas phase and in acetonitrile solution obtained from the corresponding first-principle MD trajectories at B3LYP/6-31G(d,p) level of theory

Mode	Gas phase	Solution	Δ	Exp.[84]
Out of plane NH_2	421	455	+34	
Wagging out of phase NCO/NH_2	561	572	+10	
Torsion	586	600	+12	
Out of plane CH	1,022	1,008	−12	
Wagging in phase NCO/NH_2	1,042	1,068	+27	
CN stretch	1,244	12,520	+8	
In plane COH sciss	1,389	1,375	−13	(−2)
In plane NH_2 sciss	1,605	1,610	+5	(+3)[a]
CO stretch	1,758	1,721	-37	−38
CH stretch	2,822	2,852	+29	(+33)
Sym NH stretch	3,402	3,458	+56	
Asym NH stretch	3,555	3,575	+20	

[a] Solvent shift for CH_2 scissoring of formaldehyde
Experimental solvent shifts for formamide and formaldehyde (in parenthesis) are also compared

the most important discrepancy being about 30 cm^{-1} for the C–N stretching.

The trajectory simulating the formamide in acetonitrile solution samples a phase space representing the solute instantaneously adjusted to the mean field of the polarized solvent. This approximation is analogous to the Born–Oppeheimer approximation, according to which the electronic eigenstates are considered to instantaneously adjust to the nuclear configurations. In particular, the absence of explicit solvent coordinates in the Hamiltonian involves an effective decoupling between the solute vibration and the solvent dynamics. As a consequence, the vibrational analysis does not capture dynamical effects such as the broadening of the linewidths in IR spectra. In the present case (neutral solute and non-protic, weakly polar solvent), we can safely

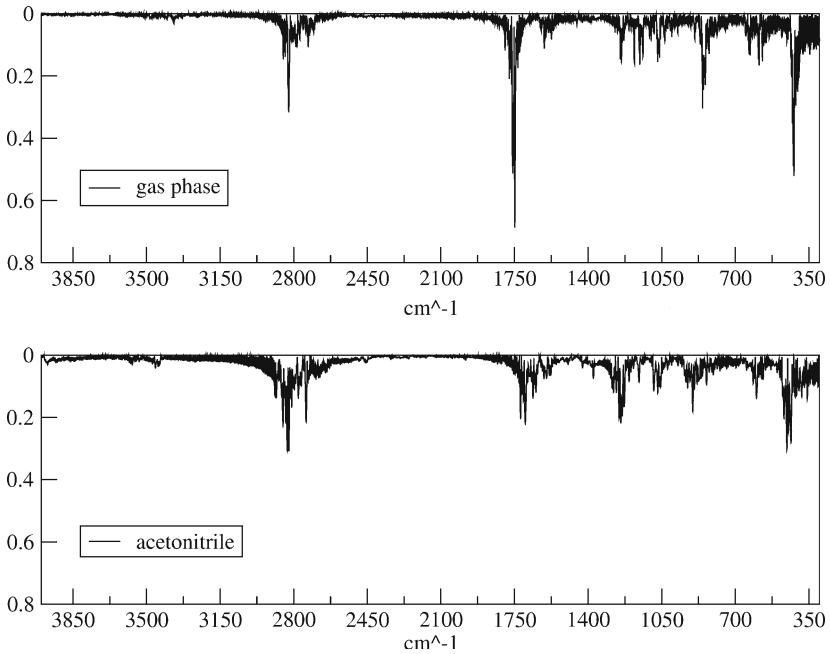

Fig. 6 IR simulated spectra for formamide in gas phase and acetonitrile solution

assume that coupling due to specific solute–solvent interactions can be neglected, while the mean field inclusion can take into account the most important features of the solvent effects on the vibrational motion, namely frequency shifts and modulation of the IR band intensities. In Table 8, we collect frequency shifts induced by acetonitrile. Corresponding values experimentally recorded for formamide and formaldehyde in the same solvent [84] are also reported for comparison. Generally, frequency variations are compatible with the increase in relative weight of the charge separation resonance structure, stabilized by the polar environment. For example, we observe a sensible red-shift of the carbonyl stretching, due to the decrease of the double bond character: the average value of the C = O bond length is 1.217 and 1.228 Å in gas phase and acetonitrile, respectively. On the other hand, the increase of the C–N double bond character is compatible with the blue-shift observed in the wagging motions, the torsion and the C–N stretching itself.

In Fig. 6, we compare IR spectra simulated both in the gas phase and acetonitrile solution. The infrared absorption coefficient, $\alpha(\omega)$, is calculated from the Fourier transform of the time correlation of the total dipole moment, of the system, $\mathbf{M}(t)$

$$\alpha(\omega)n(\omega) = \frac{2\pi\beta\omega^2}{3cV}\int \mathrm{d}t\,\langle \mathbf{M}(\mathbf{t})\mathbf{M}(\mathbf{0})\rangle \exp(i\omega t) \qquad (23)$$

where $n(\omega)$ is the refractive index of the medium at frequency ω, $\beta = 1/(kT)$, c is the speed of light in a vacuum, V is the volume and $\mathbf{M}$ is the formamide dipole moment. Brackets indicate a statistical average. Equation 23 also accounts for a quantum correction factor (multiplying the classical line shape) of the form $\beta h\omega/(1 - \exp(-\beta h\omega))$ [85].

The most evident feature is the strong reduction of the intensity of the band assigned to the carbonyl stretching. This is what one could expect due to the increasing of the polarity of the C=O bond induced by the solvent, and the consequent decreasing of the dipole transition moment.

4 Conclusions

In the present work, we have described in some detail an effective and reliable mean field based model for simulating and investigating the behavior of complex molecular systems in condensed phases, such as solute–solvent systems. The MF model allows to treat explicitly a reduced part of the system, e.g. the solute and two/three solvent shells, at different levels of theory, from very efficient, but less accurate, molecular mechanics (MM) methods to more sophisticated hybrid QM/MM or full QM methods. Interactions with the environment (solvent) are taken into account via a continuum model, which properly includes both short-range dispersion–repulsion and long-rage electrostatic contributions. Remarkably, ab initio molecular dynamics can be performed according to the Born–Oppenheimer or the extended-Lagrangian scheme allowing to exploit conveniently both post-Hartree Fock and DFT methods

using localized basis functions. In particular, popular and well-trusted DFT methods based on hybrid functionals, e.g. B3LYP and PBE0, can be used, in an efficient manner, with the ADMP methodology. Overall, the MF model can be considered a promising alternative, especially for non-periodic systems such as liquids and solutions, with respect to other standard ab initio molecular dynamics methods, like the Car–Parrinello approach.

Moreover, we have shown how accurate spectroscopic properties, such as NMR chemical shifts, UV-vis transition energies and EPR hyperfine coupling constants, of molecules in solution can be obtained by combining the present model with post-MD high-level quantum mechanical calculations. In particular, the effectiveness of such a computational strategy was demonstrated considering the case of acetone, acrolein and glycine radical in aqueous solutions and formamide in acetonitrile solution. Note that in the aqueous solution simulations part of the solvent was treated explicitly to account for the specific solute–solvent hydrogen bonding interactions, whereas acetonitrile was entirely modeled as a structureless polarizable medium. Remarkably, an overall fair agreement with experiments has been obtained in all systems considered. Specifically, it has been shown that structural fluctuations of the solute–solvent configurations during the dynamics should be properly included in the statistical averages of the computed spectroscopic properties. Besides, from the vibrational mode analysis of formamide, it has been shown that even more subtle dynamical effects can be reliably studied with the MF model.

In conclusion, we think that, at this stage, a number of physical and chemical processes in solution can be reliably explored by means of such an integrated approach, provided that the relaxation times of such processes are short enough to be accessible to standard computational facilities. At longer time scale, the use of stocastic models or other methodologies suited to enhance the configurational sampling are certainly required.

Acknowledgements The financial support of MIUR and INSTM is gratefully acknowledged. All the calculations have been performed at "Campus Computational Grid"-Università di Napoli "Federico II" advanced computing facilities.

References

1. Tomasi J, Persico M (1994) Chem Rev 94:2027
2. Cramer CJ, Truhlar DG (1999) Chem Rev 99:2161
3. Tomasi J, Mennucci B, Cammi R (2005) Chem Rev 105:2999
4. Brancato G, Nola AD, Barone V, Amadei A (2005) J Chem Phys 122:154109
5. Brancato G, Rega N, Barone V (2006) J Chem Phys 124:214505
6. Rega N, Brancato G, Barone V (2006) Chem Phys Lett 422:367
7. Bolton K, Hase WL, Peslherbe GH (1998) Modern methods for multidimensional dynamics computation in chemistry. In: Direct dynamics of reactive systems. World Scientific, Singapore, p. 143
8. Millam JM, Bakken V, Chen W, Hase WL, Schlegel HB (1999) J Chem Phys 111:3800
9. Marx D, Hutter J (2000) Modern methods and algorithms of quantum chemistry. Ab initio molecular dynamics: theory and implementation vol 1. John vonNeumann Institute for Computing, Julich p 301
10. Schlegel HB, Millam JM, Iyengar SS, Voth GA, Daniels AD, Scuseria GE, Frisch MJ (2001) J Chem Phys 114: 9758
11. Iyengar SS, Schlegel HB, Millam JM, Voth GA, Scuseria GE, Frisch MJ (2001) J Chem Phys 115:10291
12. Schlegel HB, Iyengar SS, Li X, Millam JM, Voth GA, Scuseria GE, Frisch MJ (2002) J Chem Phys 117:8694
13. Becke AD (1993) J Chem Phys 98:5648
14. Ben-Naim A (1987) Solvation thermodynamics. Plenum Press, New York
15. Orozco M, Luque FJ (2000) Chem Rev 100:4187
16. Roux B, Simonson T (1999) Biophys Chemi 78:1
17. Pascual-Ahuir JL, Silla E, Tũnon I (1994) J Comput Chem 15:1127
18. Scalmani G, Rega N, Cossi M, Barone V (2002) J Comput Meth Sci Eng 2:469
19. Pierotti RA (1976) Chem Rev 76:717
20. Claverie P (1978) Intermolecular interactions: from diatomics to biomolecules. In: Pullman B (ed). Wiley, Chichester
21. Benzi C, Cossi M, Improta R, Barone V (2005) J Comput Chem 26:1096
22. Frisch MJ, Trucks GW, Schlegel HB et al (2003) Gaussian 03, revision c.02. Gaussian, Inc.
23. Jackson JD (1999) Classical electrodynamics, 3rd edn. Wiley New York
24. Warwicker J, Watson HC (1982) J Mol Biol 157:671
25. Klapper I, Hagstrom R, Fine R, Sharp K, Honig B (1986) Proteins 1:47
26. Miertuš S, Scrocco E, Tomasi J (1981) Chem Phys 55:117
27. Zauhar R, Morgan R (1985) J Mol Biol 186:815
28. Shaw P (1985) Phys Rev A 32:2476
29. Vorobjev YN, Scheraga HA (1997) J Comp Chem 18:569
30. Benzi C, Cossi M, Barone V, Tarroni R, Zannoni C (2005) J Phys Chem B 109:2584
31. Mennucci B, Tomasi J (1997) J Chem Phys 106:5151
32. Chipman DM (1996) J Chem Phys 104:3276
33. Chipman DM (1997) J Chem Phys 106:10194
34. Zhan C-G, Chipman DM (1999) J Chem Phys 110:1611
35. Chipman DM (1999) J Chem Phys 110:8012
36. Chipman DM (2000) J Chem Phys 112:5558
37. Cossi M, Rega N, Scalmani G, Barone V (2001) J Chem Phys 114:5691
38. Klamt A, Schüurmann G (1993) J Chem Soc Perkin 2 Trans :799
39. Barone V, Cossi M (1998) J Phys Chem A 102:1995
40. Cossi M, Rega N, Scalmani G, Barone V (2003) J Comp Chem 24:669
41. Cossi M, Scalmani G, Rega N, Barone V (2002) J Chem Phys 117:43
42. Floris FM, Tomasi J (1986) J Comput Chem 10:616
43. Floris FM, Tomasi J, Pascual-Ahuir JL (1991) J Comput Chem 12:784
44. Benzi C, Crescenzi O, Pavone M, Barone V (2004) Magn Reson Chem 42:S57

45. Gustavsson T, Banyasz A, Lazzarotto E, Markovitsi D, Scalmani G, Frisch MJ, Barone V, Improta R (2006) J Am Chem Soc 128:607
46. Barone V, Improta R (2004) Chem Rev 104:1231
47. Pavone M, Cimino P, De Angelis F, Barone V (2006) J Am Chem Soc 128:4338
48. Brancato G, Rega N, Barone V (2006) J Chem Phys 125:164515
49. Pavone M, Brancato G, Morelli G, Barone V (2006) Chem Phys Chem 7:148
50. Cossi M, Crescenzi O (2003) J Chem Phys 118:8863
51. Improta R, Barone V, Kudin KN, Scuseria GE (2002) J Am Chem Soc 124:113
52. Doltsinis NL, Sprik M (2000) Chem Phys Lett 330:563
53. Car R, Parrinello M (1985) Phys Rev Lett 55:2471
54. Stephens PJ, Devlin FJ, Chabalowski CF, Frisch MJ (1994) J Phys Chem 98:11623
55. Freindorf M, Shao Y, Furlani TR, Kong J (2005) J Comput Chem 26:1270
56. Jameson CJ (1996) Ann Rev Phys Chem 47:135
57. Korzhnev DM, Billeter M, Arseniev AS, Orekhov VY (2001) Prog Nucl Magn Reson Spectrosc 38:197
58. Helgaker T, Jaszunski M, Ruud K (1999) Chem Rev 99:293
59. Wolinski K, Hilton JF, Pulay P (1990) J Am Chem Soc 112:8251
60. Cheeseman JR, Trucks GW, Keith TA, Frisch MJ (1996) J Chem Phys 104:5497
61. Bayliss NS, McRae EG (1954) J Phys Chem 58:1006
62. Balasubramanian A, Rao CNR (1962) Spectrochim Acta 18:1337
63. Bayliss NS, Wills-Johnson G (1968) Spectrochim ActaA 24:551
64. Blair JT, Krogh-Jespersen K, Levy RM (1989) J Am Chem Soc 111:6948
65. Kongsted J, Osted A, Pedersen TB, Mikkelsen KV, Christiansen O (2004) J Phys Chem A 108:8624
66. Bayliss NS, McRae EG (1954) J Phys Chem 58:1002
67. Crescenzi O, Pavone M, De Angelis F, Barone V (2005) J Phys Chem B 109:445
68. Buswell AM, Dunlop EC, Rodebush WH, Swartz JB (1940) J Am Chem Soc 62:325
69. Walsh AD (1945) Trans Faraday Soc 41:498
70. Mackinney G, Temmer O (1948) J Am Chem Soc 70:3586
71. Inuzuka K (1961) Bull Chem Soc Jpn 34:6
72. Becker RS, Inuzuka K, King J (1970) J Chem Phys 52:5164
73. Bair EJ, Goietz W, Ramsay DA (1971) Can J Chem 49:2710
74. Osborne GA, Ramsay DA (1973) Can J Chem 51:1170
75. Luthy A (1923) Z Phys Chem 107:284
76. Blacet FE, Young WG, Roof JG (1937) J Am Chem Soc 59:608
77. Inuzuka K (1960) Bull Chem Soc Jpn 33:678
78. von Paul H, Fisher H (1971) Helv Chim Acta 54:485
79. Neta P, Fessenden RW (1971) J Phys Chem 75:738
80. Ghosh DK, Wiffen DH (1960) J Chem Soc :1869
81. Rega N, Cossi M, Barone V (1997) J Am Chem Soc 119:12962
82. Barone V (2004) J Phys Chem A 108:4146
83. Barone V (2005) J Chem Phys 122:014108
84. Wong MW, Wiberg KB, Frisch MJ (1991) J Chem Phys 95:8991
85. Ramirez R, Lopez-Ciudad T, Kumar P, Marx D (2004) J Chem Phys 121:3973