Spectroscopic Properties in the Liquid Phase: Combining High-Level Ab Initio Calculations and Classical Molecular Dynamics

ARTICLE *in* CHEMPHYSCHEM · JANUARY 2006

Impact Factor: 3.42 · DOI: 10.1002/cphc.200500357 · Source: PubMed

4 AUTHORS, INCLUDING:

Michele Pavone
University of Naples Federico II
58 PUBLICATIONS **1,183** CITATIONS
SEE PROFILE

Vincenzo Barone
Scuola Normale Superiore di Pisa
773 PUBLICATIONS **44,695** CITATIONS
SEE PROFILE

CHEMPHYSCHEM

DOI: 10.1002/cphc.200500357

Spectroscopic Properties in the Liquid Phase: Combining High-Level Ab Initio Calculations and Classical Molecular Dynamics

Michele Pavone, Giuseppe Brancato, Giovanni Morelli, and Vincenzo Barone*[a]

We present an integrated computational tool, rooted in density functional theory, the polarizable continuum model, and classical molecular dynamics employing spherical boundary conditions, to study the spectroscopic observables of molecules in solution. As a test case, a modified OPLS-AA force field has been developed and used to compute the UV and NMR spectra of acetone in aqueous solution. The results show that provided the classical force fields are carefully reparameterized and validated, the proposed approach is robust and effective, and can also be used by nonspecialists to provide a general and powerful complement to experimental techniques.

1. Introduction

Recent advances in the reliability and effectiveness of quantum mechanical methods, and their implementation in popular software packages, have provided a large community of chemists with sophisticated tools that allow the theoretical reproduction and interpretation of several spectroscopic properties.[1–7] Also, the diffusion of fast parallel computers and the development of computationally efficient linear scaling methods[2] make possible the use of first-principle calculations on complex molecular systems, such as large biomolecules or synthetic polymers. As a consequence, a closer collaboration between computational chemists and experimentalists is emerging. However, despite the remarkable success of quantum chemistry, the accurate computation of spectroscopic observables is generally challenging. In fact, the vast majority of experimental data, for example UV excitation energies or NMR chemical shifts, are recorded in solution, which represents the natural physical environment for many chemical systems ranging from living cells to industrial catalysts. In such conditions, an appropriate theoretical model must take into account many subtle effects.[8–16]

Herein, we consider the typical case of a solute/solvent system, where the solute molecule is the spectroscopically active moiety, which represents a large class of molecular systems in the condensed phase. For such systems, the physical response of the solute to an applied electromagnetic field is generally quite a complex phenomenon, which is affected by both the interaction with the environment, for example the solvent, and the fluctuations of the solute geometry as well as the surrounding solvent molecules. From a theoretical point of view, the problem of reproducing accurately spectroscopic parameters of molecules in solution can be decomposed into two parts: 1) sampling a representative region of the configurational space of the solute/solvent system; and 2) performing spectroscopic calculations, at a high level of theory, on the sampled molecular structures and then obtaining the physical observables from statistical averages.

The latter point can be effectively addressed by recently developed quantum mechanical (QM) methods based on post-Hartree–Fock (HF) or density functional theory (DFT) calculations, especially when second-order properties are considered.[4,6] In such methods the environmental effects on the solute molecule are usually taken into account by including explicitly the solvent molecules in the QM treatment[9] or, more efficiently, by embedding the solute in a dielectric continuum that accounts for the solvent polarization.[11,12] Also, mixed quantum/classical (QM/MM) methods are often used to treat explicitly extended molecular systems at low computational cost, by reducing the expensive ab initio calculations to the chemically important region of the system.[15,16]

Less straightforward is the determination of an accurate and statistically significant sampling of the solute/solvent system, which is, in most cases, unavoidable due to the sensitivity of spectroscopic properties to the specific molecular configurations. Ab initio molecular dynamics (MD) techniques, such as the popular Car–Parrinello MD, are usually the methods of choice when high accuracy and reliability are required.[17] Unfortunately, ab initio MD simulations have some severe limitations due to the high computational cost: typically, system sizes are limited to a few hundreds of atoms and simulation lengths are of the order of tens of picoseconds. Moreover, the usual QM models adopted in ab initio MD, for example, DFT with functionals based on the generalized gradient approxima-

[a] *Dr. M. Pavone, Dr. G. Brancato, Dr. G. Morelli, Prof. V. Barone*
Dipartimento di Chimica, Università di Napoli "Federico II"
Complesso Universitario di Monte Sant'Angelo
Via Cintia, 80126 Napoli (Italy)
Fax: (+39) 081-674-090
E-mail: baronev@unina.it

tion (GGA), may not be accurate enough for reproducing geometrical and energetic features of molecular systems. As a consequence, the computed spectroscopic properties significantly deviate from experiments (see the Results and Discussion section).[16,18,19] On the other hand, classical MD methods are computationally very efficient but their reliability is subject to the implicit approximations included in the empirically derived force fields. Nevertheless, classical MD is well-suited to computing statistical mechanical properties from extended simulations.

As a result of an ongoing effort to develop and validate a general protocol for the QM computation of spectroscopic parameters for large molecules in solution, here we propose an effective and flexible strategy based on both ab initio and classical methods. Although this kind of approach has quite a long tradition,[15,16,18,20] there are a number of specific issues that require further consideration. Our specific strategy consists of a three-step procedure:

1. The parameterization of a reliable force field for the solute/solvent system to be used in long classical MD simulations. Parameters can be obtained from existing molecular force fields, such as OPLS,[21] AMBER,[22] or CHARMM,[23] and also derived, if necessary, from more sophisticated QM calculations and dynamics. In particular, intermolecular potentials can be conveniently adjusted to reproduce solute/solvent radial distribution functions as obtained from short ab initio MD simulations.
2. The extended sampling of the molecular system via classical MD simulations, whose force field has been previously obtained and tested, using the mean field (MF) method.[24] Briefly, the MF is a continuum model suited for simulation of solute/solvent systems: the solute molecule is placed at the center of a spherical cavity embedded in a dielectric continuum and solvated with a large number (hundreds or thousands) of explicit solvent molecules.
3. The QM computation of the spectroscopic parameters. Remarkably, for each sampled configuration, the whole molecular system is considered in such calculations, which also include the dielectric continuum for consistency. Clearly, only a few solvent molecules along with the solute are treated at the quantum level, whereas the remaining molecules are included as point charges in what can be more properly named a QM/MM calculation.

In a previous publication,[24] it was shown that the MF model is very effective in reproducing the structural and thermodynamic properties of bulk water as well as aqueous solutions of simple solutes, as compared to conventional simulations using periodic boundary conditions (PBC). In this context, the use of the MF model is very well-suited because it elegantly allows retention of the same explicit molecular system, as obtained from the simulated trajectory (step 2) in the QM calculations (step 3). Moreover, the MF model can be conveniently used for QM or mixed QM/MM molecular dynamics simulations of liquids and solutions by allowing the use of localized basis sets instead of plane waves, as is usual in condensed-phase simulations. Notably, our choice to compute the molecular spectroscopic observables "a posteriori", at the best QM level according to the specific property considered, is to be regarded as a methodologically different approach with respect to other procedures in which the spectroscopic parameters are computed "on the fly"[25] or by using perturbative theoretical models.[26] Finally, within the procedure depicted above, a number of strictly system-dependent technical issues have to be addressed, such as the force-field parameterization and the convergence of the spectroscopic properties with the number of configurations. In particular, special care is due to all those features of the molecular force field that highly affect the final results of the computed spectroscopic properties, and not just the geometrical parameters.

Herein, a theoretical investigation of the $n\rightarrow\pi^*$ electronic transition energies and the ^{13}C and ^{17}O nuclear magnetic shielding constants of the carbonyl group of the acetone molecule in both the gas phase and aqueous solution is presented as a test of the proposed approach. Of course, a more crucial test would be provided by a completely hydrogen-bonded molecule that offered donor and acceptor possibilities to the solvent. However, acetone has become a benchmark[27] and, contrary to expectations, its spectroscopic properties are quite dependent on the details of both intra- and intermolecular interactions (see below).

First, a classical force field was derived for the intra- and intermolecular potentials of acetone from high-level ab initio calculations and Car–Parrinello molecular dynamics. Then, a large sampling of both the isolated molecule and the solute/solvent configurational space was obtained from MD simulations of about 1 ns each, from which the spectroscopic observables were computed with high accuracy. Previously, some theoretical studies of the acetone/water system have been reported using a full QM level[18,28] and a QM/MM approach,[16] in which quite short dynamics of a few picoseconds were used to compute electronic transition energies. As a consequence, the results were affected by a statistical error which was about an order of magnitude larger than in the present study.

This article is organized as follows. In Section 2, we describe the theoretical and computational details of all the calculations and simulations performed in the three stages of the proposed approach, as outlined previously. In particular, details on the crucial force-field parameterization for the acetone/water system are provided. In the Results and Discussion (Section 3), we present a convergence test concerning the number of solvent molecules to be treated at the QM level to reproduce accurately the spectroscopic parameters. Also, some basic structural and thermodynamic analyses of the MF-based MD simulation are reported, and the calculated solute spectroscopic properties are compared with those from experiments. Conclusions are given in Section 4.

2. Methods and Computational Details

2.1. Force-Field Parameters

QM calculations were carried out on acetone and the acetone-$(H_2O)_2$ cluster at different levels of theory to obtain a reliable

molecular mechanics (MM) force field for the acetone intra- and intermolecular interactions. In particular, the Perdew–Burke–Ernzerhof (PBE) functional[29] and its hybrid HF-DFT counterpart, known as PBE0,[30] were tested. As expected,[18,19,28] the inclusion of some HF exchange into the functional formulation ensured a better agreement with experimental structural parameters. Also, the results of the latest room-temperature Car–Parrinello dynamics of acetone in aqueous solution[28] were used to adjust the acetone/water intermolecular parameters, as discussed in the Results section. All QM results were obtained using the Gaussian 03 package.[31]

The final MM parameters are summarized in Table 1 and the minimum energy configuration, along with the atom-type labels, is depicted in Figure 1. In particular, the important C=O bond stretching and out-of-plane modes of acetone were fitted on QM data with Morse and harmonic potentials, respectively, whereas all other force constants were taken from the OPLS-AA force field,[21] and the equilibrium values were scaled to agree with QM-optimized structures. Note that two dummy atoms (atom label X) representing the oxygen lone pairs, which are usually absent in standard force fields, were included to account for the hydrogen-bond directionality. Partial charges of the acetone were computed by fitting an ab initio derived electrostatic potential on a grid of points around the molecule, according to the restrained electrostatic potential (RESP) procedure.[32] The water model adopted for the MM calculations of the cluster and aqueous solution is the popular simple point charge (SPC) model.[33] A modified version (see Section 2.3) of the Gromacs package[34] was used for all the MM and MD calculations.

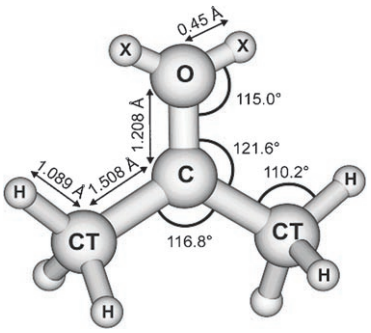

Figure 1. Structure and atom-type labeling of the acetone molecule.

Table 1. Acetone and acetone/water force-field parameters.

Bonded interactions			Nonbonded interactions		
Parameter[a]	Value	Units	Parameter	Value	Units
D (CO)	810.82	kJ mol^{-1}	q_O	0.0	
B (CO)	2.2636	Å^{-1}	q_X	−0.3023	
r_0 (CO)	1.2080[b] (1.2170[c])	Å	q_C	0.5825	
k_b (CC)	2652.66	kJ mol^{-1} Å^{-2}	q_{CT}	−0.3204	
r_0 (CC)	1.5088	Å	q_H	0.1105	
k_b (CH)	2845.12	kJ mol^{-1} Å^{-2}	$\varepsilon_{O,H}$	0.4697	kJ mol^{-1}
r_0 (CH)	1.0895	Å	$\sigma_{O,H}$	2.9500	Å
k_θ (OCC)	669.44	kJ mol^{-1} rad^{-2}	$\varepsilon_{CT,H}$	0.09310	kJ mol^{-1}
θ_0 (OCC)	121.79	deg	$\sigma_{CT,H}$	3.0000	Å
k_θ (CCC)	527.18	kJ mol^{-1} rad^{-2}	$\varepsilon_{H,H}$	0.12552	kJ mol^{-1}
θ_0 (CCC)	116.43	deg	$\sigma_{H,H}$	2.5000	Å
k_θ (HCC)	418.40	kJ mol^{-1} rad^{-2}	$\varepsilon_{Ow,O}$	0.76427	kJ mol^{-1}
θ_0 (HCC)	110.26	deg	$\sigma_{Ow,O}$	3.2980	Å
k_θ (HCH)	292.88	kJ mol^{-1} rad^{-2}	$\varepsilon_{Ow,C}$	0.53438	kJ mol^{-1}
θ_0 (HCH)	108.62	deg	$\sigma_{Ow,C}$	3.4580	Å
k_ξ	115.0	kJ mol^{-1} rad^{-2}	$\varepsilon_{Ow,CT}$	0.42367	kJ mol^{-1}
ξ_0	0.0	deg	$\sigma_{Ow,CT}$	3.3330	Å
k_ϕ (OCCH)	1.6736	kJ mol^{-1}	$\sigma_{Ow,H}$	0.28564	kJ mol^{-1}
n (OCCH)	1		$\sigma_{Ow,H}$	2.5830	Å
ϕ_0 (OCCH)	0.0	deg			
k_ϕ (OCCH)	0.16736	kJ mol^{-1}			
n (OCCH)	3				
ϕ_0 (OCCH)	180.0	deg			

[a] ξ is the OCCC improper dihedral angle. [b] Gas-phase parameter. [c] Condensed-phase parameter (see Section 3.1).

2.2. Classical Molecular Dynamics

Classical MD simulations of acetone in the gas phase and in aqueous solution were performed in a canonical ensemble at 300 K. In both cases, the overall simulation length was 1 ns, including an initial equilibration of 400 ps. The condensed-phase MD simulation was carried out using the MF model[24] with spherical boundary conditions, a novel MD approach based on a continuum model, which was recently introduced by our group. The details of the model are fully described in ref. [24]. Briefly, it consists of simulating explicitly a molecular system, such as a liquid or a solution, inside a cavity of a continuum, which accounts for the long-range interactions with the environment, thus avoiding the use of periodic boundary conditions that, in principle, could introduce unwanted correlation effects. In particular, the MF model includes both an electrostatic "reaction field" and a dispersion–repulsion contribution. In this work, the acetone molecule, solvated with 1111 SPC water molecules [$(N_{water}+1)/V = 33.184$ nm^{-3}], was placed at the center of a spherical cavity embedded in the continuum medium. The roto-translational motions of the solute were removed during the dynamics using the method described in ref. [35]. All simulations were performed in double precision with a modified version of the GROMACS simulation package,[34] which includes the MF model and the Gaussian isokinetic thermostat.[36]

2.3. Quantum Mechanical Calculations of Spectroscopic Properties

All the spectroscopic properties were computed using the Gaussian 03 package.[31] Excitation energies were computed within the TD-DFT[4] formalism employ-

ing the PBE0[30] functional and a basis set that combines the 6-311++G(2d,2p) basis on the carbonyl atoms and the 6-31+G(d,p) basis set on the other acetone and water atoms (henceforth denoted as [6-311++G(2d,2p), 6-31+G(d,p)]).[37] The consistency of this basis set has already been proven in other recent work, in which it was successfully applied to the computation of UV parameters.[27] The same basis set was also employed for the NMR shielding tensor calculations within the gauge-including atomic orbital (GIAO) formalism[5] at the PBE0[30] and MP2[38] levels of theory. The spectroscopic data were computed on different frames extracted from the gas phase and from the aqueous solution dynamics, after testing the sampling set to achieve a consistent number of snapshots and ensure converged average parameters. The solvent was taken into account within the mean-field approach, in close agreement with the MD simulations. Water molecules embedded in the continuum sphere, modeling the polarizable medium, were treated as point charges and the closest water molecules to the carbonyl oxygen were included together with the acetone molecule in the QM layer. The charges on the water atoms are those of the SPC model,[33] and both the sampling set and the number of water molecules to be included in the QM region were tested.

3. Results and Discussion

3.1. Force-Field Parameterization

The gas-phase structure of the acetone molecule is known experimentally from electron diffraction spectroscopy,[39] and was used as a reference to test the accuracy of the ab initio calculations as well as the molecular mechanics force field. First, we carried out QM calculations of acetone at different levels of theory to determine the optimal combination of ab initio methodology and basis set. To this end, both the PBE functional[29] and its hybrid HF-DFT counterpart, known as PBE0,[30] were tested. Also, the consistency of the structural parameters computed with the PBE0 functional was checked with different basis sets, from the double and triple ζ of the series of Pople et al.[37] to the correlation-consistent triple ζ of Dunning's series.[40] The acetone structural parameters obtained from experiments and computations are listed in Table 2.

A significant difference in the computed values between PBE and PBE0 with the same 6-31+G(d,p) basis set is observed: the hybrid functional increases the accuracy of the bond distances, whereas the angle values are not affected. Among the basis sets considered with the PBE0 functional, the triple-ζ set of the series of Pople et al., with diffuse and d-polarization functions on carbon and oxygen atoms and p-polarization functions on hydrogen atoms, provides geometrical parameters which are consistent with much larger basis sets, such as cc-pVTZ and Aug-cc-pVTZ, and are in close agreement with experiments, with the exception of the C–H bond. Note that the experimental value of the latter is affected by a larger error and all computational results reported herein are consistent. After all these tests we found that PBE0/6-311+G(d,p) represents the best compromise between accuracy and efficiency, and all QM calculations used in the force-field parameterization were performed at this level of theory.

The acetone MM structure is also listed in Table 2: a very good agreement with both high-level QM calculations and experimental data was obtained. As mentioned in Section 2.1, PBE0/6-311+G(d,p) calculations were performed to fit a Morse potential for the C=O bond stretching [Eq. (1)]:

$$V_{\mathrm{bond}}(r) = D[1 - \exp(-B(r - r_0))]^2 \quad (1)$$

where D, B, and r_0 are the usual Morse parameters. This functional form was chosen, instead of the usual harmonic potential, to account for the anharmonic character of the bond. Also, the harmonic dihedral parameters that correspond to the out-of-plane mode of the acetone were fitted at the same level of theory, thus ensuring a reliable force constant [Eq. (2)]:

$$V_{\mathrm{dihedral}}(\xi) = k_{\xi}(\xi - \xi_0)^2 \quad (2)$$

where ξ is the dihedral angle and k_{ξ} the corresponding force constant (see Table 1 for parameter values). The correct sampling of these two modes during the dynamics was found to be crucial for accurate calculation of the spectroscopic properties. Moreover, lone pairs were added to the oxygen atom of the acetone molecule to reproduce the QM minimum-energy structure of the acetone-$(H_2O)_2$ cluster: the correct hydrogen-bond geometry, as shown in Table 3, was obtained with an O···X distance of 0.45 Å and a C=O–X angle of 115° (see Figure 1). Note that commonly used force fields, such as OPLS, CHARMM, and GROMOS, do not include lone pairs attached to the carbonyl oxygen atom and therefore are unable to provide some basic features of the hydrogen-bond pattern (see next section). Also, it was necessary to impose a different equilibrium distance, r_0, for the C=O bond in the condensed-phase simulation with respect to the gas phase ($r_{0,\mathrm{gas\,phase}} = 1.2080$ Å; $r_{0,\mathrm{cond.\,phase}} = 1.2170$ Å) to account for the elongation of the carbonyl group, as observed in Table 3. The C=O bond stretching is a result of the polarization effect of the solvent, which induces a change of the bond order from

Table 2. Geometrical parameters of acetone optimized at different computational levels are compared with their experimental counterparts. Bond distances are in Å, angles in degrees.

	C=O	C–C	C–H	C–C–C	C–C=O
PBE/6-31+G(d,p)	1.229	1.520	1.102	116.6	121.7
PBE0/6-31+G(d,p)	1.214	1.510	1.094	116.6	121.7
PBE0/6-311+G(d,p)	1.208	1.509	1.090	116.4	121.8
PBE0/cc-pVTZ	1.206	1.507	1.091	116.4	121.8
PBE0/Aug-cc-pVTZ	1.207	1.506	1.091	116.4	121.8
MM	1.208	1.508	1.089	116.8	121.6
experiment[39]	1.210 (0.003)	1.507 (0.002)	1.076 (0.006)	116.7 (0.3)	121.7[a]

[a] Assumed value.

Table 3. Geometric parameters of the acetone-$(H_2O)_2$ cluster arising from QM and MM geometry optimizations. QM computations were performed at the PBE0/6-311+G(d,p) level, and MM calculations using the parameters of Table 1. Water atoms are indicated by the subscript w. Bond distances are in Å, angles in degrees.

	MM	QM	QM+PCM
O···H_w	1.91	1.91	1.88
O···O_w	2.91	2.85	2.85
C=O···H_w	115.4	117.2	126.4
O···O_w–H_w	1.8	12.2	0.2
C=O	1.211[a] (1.220[b])	1.221	1.224

[a] Value obtained with r_0=1.2080 Å. [b] Value obtained with r_0=1.2170 Å.

a purely double-bond character to a mixed double/single bond. Such a structural modification ultimately depends on an electronic density rearrangement and cannot be trivially obtained with an effective force field. For such a reason, we decided to use two different parameters for the MM equilibrium distance of the carbonyl group (see Table 1), according to high-level ab initio molecular geometries. Furthermore, it is interesting to note that solvent has a negligible effect on the C=O bond distance beyond the two closest water molecules (≈0.003 Å), as shown by the inclusion of a solvent effect in the geometry optimization with the polarizable continuum model (PCM; see last column of Table 3), which is also consistent with the MM results (for the acetone-$(H_2O)_2$ cluster: $r_{C=O}$= 1.220 Å; for acetone in solution: $r_{C=O}$=1.222 Å; see Table 4).

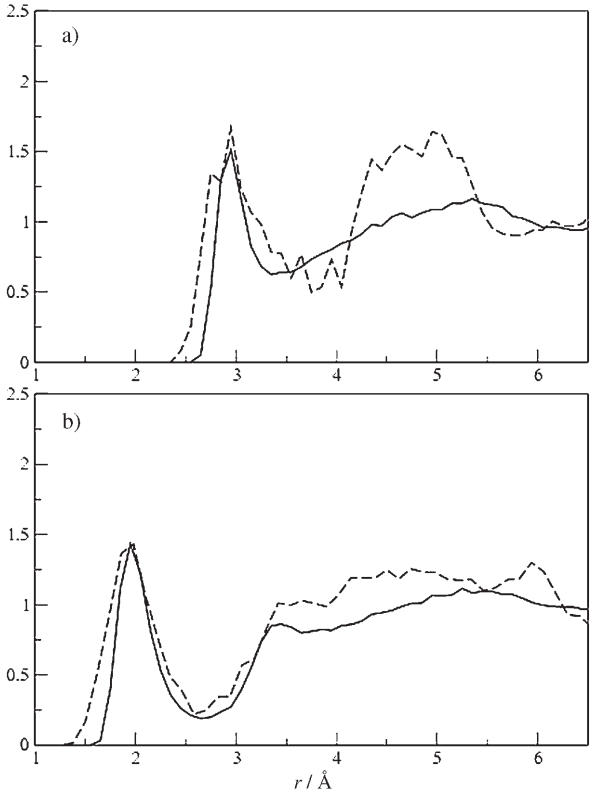

Figure 2. a) $O_{acetone}$–O_{water} and b) $O_{acetone}$–H_{water} radial distribution functions. MM molecular dynamics (——); Car–Parrinello molecular dynamics (-----) from ref. [26].

Table 4. Average geometrical parameters of acetone issuing from the gas phase and aqueous solution dynamics. Bond distances are in Å, angles in degrees, and standard deviations in parentheses.

	C=O	C–C	C–H	C–C–C	C–C=O
gas-phase MM[a]	1.209 (0.002)	1.514 (0.003)	1.089 (0.002)	116.9 (1.1)	121.4 (0.6)
gas-phase CPMD	1.23 (0.02)	1.52 (0.04)	1.11 (0.04)	116.9 (5.3)	121.4 (4.3)
solution MM[b]	1.222 (0.002)	1.511 (0.002)	1.090 (0.001)	115.7 (0.6)	121.5 (0.5)
solution CPMD	1.25 (0.01)	1.50 (0.03)	1.11 (0.02)	117.4 (3.1)	121.0 (3.4)

[a] Values obtained with r_0=1.2080 Å. [b] Values obtained with r_0=1.2170 Å (see Table 1).

Finally, the acetone/water intermolecular parameters, for example, the Lennard–Jones interaction between the oxygen atoms, were adjusted to allow the structural properties to agree with our previous condensed-phase CPMD simulation[28] (see Figure 2), as discussed below.

3.2. Molecular Dynamics Simulations

Constant-temperature MD simulations of acetone at 300 K in both vacuum and aqueous solution were carried out to produce a large sampling of the configurational space. Here, we compare the results of the classical (MM) dynamics with the ab initio MD simulations (CPMD), based on the Car–Parrinello scheme and the PBE density functional of ref. [28]. Note that, as usual, CPMD simulations are very short (about 5.5 ps) and, in some cases, only a qualitative comparison with MM results is possible.

In Table 4, some average geometrical parameters of acetone are reported. Overall the molecular structure has not changed significantly in going from the gas phase to liquid: only the C=O bond is slightly elongated, as a result of the interactions with the solvent. The main difference between CPMD and MM values consists of the C=O average distance. As mentioned in the Methods section, the MM carbonyl group was parameterized with an improved DFT functional (PBE0) with respect to CPMD.

In the condensed phase, the hydrogen-bond arrangement is very similar in both ab initio and classical simulations, as shown by different structural analyses. The $O_{acetone}$–O_{water} and $O_{acetone}$–H_{water} radial distribution functions (RDFs) are comparable (see Figure 2): the first peaks show remarkably the same positions and heights, whereas a significant deviation is observed in the RDFs at a larger distance. However, CPMD results are clearly too noisy due to the poor statistics and a detailed comparison seems meaningless. Somewhat more reliable are the results of the average number of hydrogen bonds (identified using the same criteria as in ref. [28]) formed by acetone

and water, the corresponding distribution of one, two, and three hydrogen bonds, and the average $O_{acetone}$–H_{water} distance. As reported in Table 5, the MM and CPMD results agree nicely.

Table 5. Main characteristics of acetone–water hydrogen bonds.

	CPMD[28]	MM
number of hydrogen bonds	2.1	2.0
% of one hydrogen bond	17	15
% of two hydrogen bonds	52	66
% of three hydrogen bonds	31	19
O···H_w	1.99 Å	2.06 Å

Note that, in agreement with CPMD, a non-negligible number of configurations in the MM simulation show one or three hydrogen bonds, with a slight preference for the latter.

Furthermore, we want to stress once more the importance of the inclusion of the lone pairs on the oxygen atom of the acetone molecule. In Figure 3, the structural effect of the lone

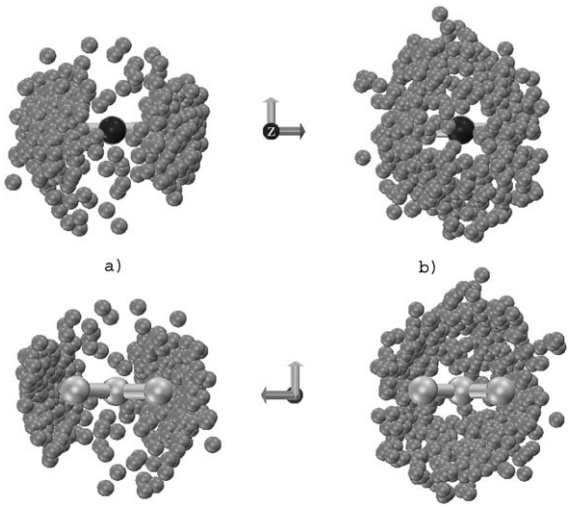

Figure 3. Distribution of the water molecules that form hydrogen bonds with acetone. a) Acetone with lone pairs on the oxygen atom: acetone oxygen atom, black; carbon, light grey; water oxygen atoms, grey. b) Acetone without lone pairs on the oxygen atom. Hydrogen atoms are omitted for clarity.

pairs on the angular distribution of the hydrogen bonds is depicted. As expected, water molecules preferentially lie close to the acetone molecular plane (Figure 3a), whereas the absence of lone pairs generates a spherically symmetric distribution around the carbonyl group (Figure 3b).

3.3. Spectroscopic Parameters

In this section, we compare the solvent shifts of the $n\rightarrow\pi^*$ electronic transition energy and the ^{13}C and ^{17}O nuclear magnetic shieldings for the passage from the gas phase to aqueous solution obtained by high-level QM computations with their experimental counterparts.[41–43] According to the methodological approach proposed above, we extracted a large number of molecular configurations from the simulated trajectories and obtained the spectroscopic observables as averages of the corresponding computed parameters on each configuration. Depending on the molecular system considered and on the specific spectroscopic property, some technical issues should be addressed, such as the choice of the QM method, the size of the QM region in the QM/MM calculations, and the number of configurations required to reach a convergence of the computed spectroscopic values. As mentioned in the Methods section, UV transition energies were evaluated with the TD-DFT approach and NMR shieldings by GIAO-DFT calculations using PBE0/[6-311++G(2d,2p), 6-31+G(d,p)] in both cases. NMR parameters were computed at the MP2 level for comparison.

Some tests were carried out to check the convergence of spectroscopic parameters with the number of water molecules treated at the QM level along with the acetone, while the remaining solvent molecules were included in the Hamiltonian as point charges according to the SPC model (MM), as described in Section 2.4. In Tables 6 and 7, we report the results for the NMR ^{17}O shieldings and the $n\rightarrow\pi^*$ excitation energy, respectively, computed with 0, 2, 3, and 5 QM water molecules. Also, in Table 6 we have included for comparison the results obtained by applying the PCM model directly on the QM

Table 6. Convergence of NMR parameters as a function of the number of explicit water molecules included in the QM layer. ^{17}O isotropic shielding constants σ_{iso} [ppm] were computed at the PBE0/[6-311++G(2d,2p), 6-31+G(d,p)] level of theory; MM waters are the SPCs.

	σ_{iso}[a]	σ_{iso}[b]
0 QM+MM	−258	−249
2 QM+MM	−253	−243
3 QM+MM	−253	−244
5 QM+MM	−252	−241
0 QM+PCM	−299	−289
2 QM+PCM	−261	−243
3 QM+PCM	−258	−245
0 QM	−362	−349
2 QM	–	−283
3 QM	–	−280
5 QM	–	−260
35 QM	–	−235
55 QM	–	−239

[a] Average values computed on 100 configurations. [b] Values computed on one specific configuration.

Table 7. n→π* excitation energies computed at the PBE0/[6-311++G(2d,2p), 6-31+G(d,p)] level as a function of the number of explicit water molecules included in the QM layer. MM waters are represented by SPCs and average values are computed over 100 configurations.

	$\Delta E_{(n-\pi^*)}$ [eV]
0 QM+MM	4.56
2 QM+MM	4.57
3 QM+MM	4.57
5 QM+MM	4.57

region (xQM+PCM), that is, avoiding the inclusion of any point charges, and the values obtained for clusters of different size without the continuum (xQM). In particular, the slow convergence with cluster size of the computed NMR shieldings, as observed for one particular molecular configuration (see the xQM results in last column of Table 6), clearly demonstrates the importance of the inclusion of long-range bulk solvent effects, either by continuum models (e.g., PCM) or by explicit point charges (MM). Moreover, the full QM result for the acetone-$(H_2O)_{55}$ cluster, isotropic shielding constant $\sigma_{iso}=-239$ ppm, computed on a specific molecular configuration, is very close to the result obtained considering only five water molecules in the QM region, $\sigma_{iso}=-244$ ppm, for the same cluster (5 QM+50 MM). Remarkably, the two approaches investigated that include solvent effects (xQM+PCM and xQM+MM) do converge to values in close agreement with each other (a difference of about 5 ppm should be considered as within statistical error), and both are computationally very efficient for a given size of the QM region. It should be pointed out that, in general, a very limited number of water molecules, typically those which form hydrogen bonds with the carbonyl group, have to be included in the QM region to obtain an acceptable accuracy, as noted in our previous paper.[28] Here, we have found that UV energies are not very much affected by the number of QM water molecules (see Table 7), whereas the NMR results are unchanged only with two or more QM molecules (see Table 6). In the latter case, the inclusion of quantum effects in the treatment of solute–solvent interactions is crucial, so that perturbative approaches,[26] which account only for the electrostatic contribution of the environment, are expected to be inadequate. In the following, according to the methodology proposed herein, we always performed spectroscopic calculations with the two water molecules closest to the carbonyl oxygen atom treated at the QM level and the remaining water molecules treated as point charges plus a reaction field.

The accuracy of the spectroscopic observables was tested against the number of uncorrelated molecular configurations obtained from the MD simulations. As shown in Figure 4, a few tens of configurations represent a statistically significant data set for an accurate estimate of the spectroscopic parameters, both in the gas phase and in solution. Due to the low computational cost of such calculations, the results reported below are obtained from averages over 100 molecular configurations taken from the simulated trajectories at evenly spaced time intervals.

The ^{13}C and ^{17}O isotropic shielding constants relative to the corresponding carbonyl atoms of acetone are reported in Table 8, as obtained from the acetone gas phase and aqueous solution molecular dynamics and computed at both the PBE0 and MP2 levels of theory. The change in the ^{13}C NMR shieldings on going from the gas phase to liquid shows no differences between the PBE0 and MP2 results and, notably, is in excellent agreement with experiment. On the other hand, the calculated ^{17}O shieldings are somewhat less accurate: as shown in Table 8, PBE0 overestimates the solvent effects on the NMR shieldings ($\sigma_{iso}=87$ ppm) with respect to experiment ($\sigma_{iso}=75.5$ ppm), whereas MP2 underestimates the same parameter ($\sigma_{iso}=65$ ppm). However, an error of about 10 ppm should be considered acceptable and it should be noted that no scaling factors or other kind of corrections, based on empirical considerations, were applied to the reported results. Also, the solvent shifts of the ^{13}C and ^{17}O shielding constants computed at the PBE0 level on the CPMD trajectories are −23.1 and 105, respectively, which are significantly different from experiment due to the overestimated C=O bond length.

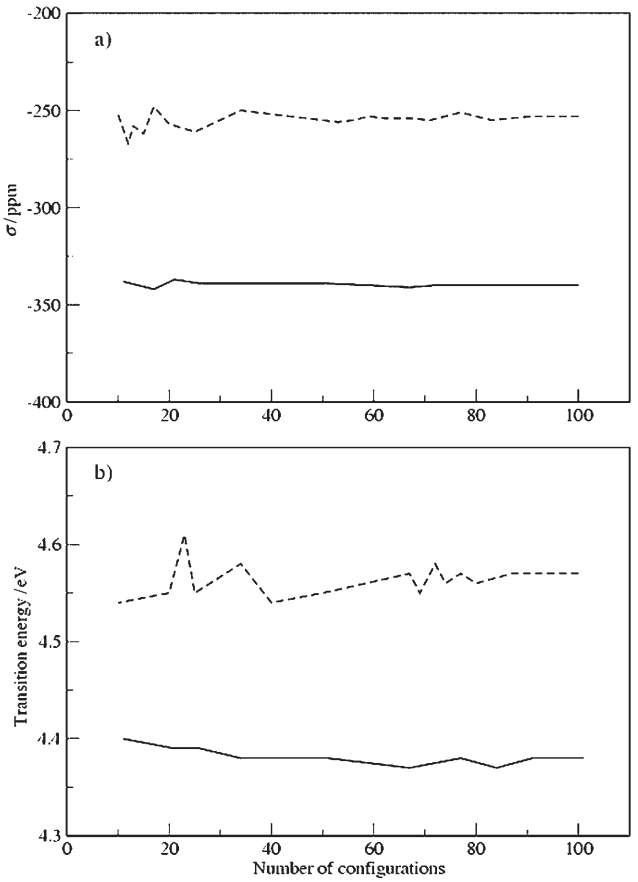

Figure 4. a) Average ^{17}O isotropic shielding constant of acetone as a function of the number of configurations. b) Average $n\rightarrow\pi^*$ excitation energy as a function of the number of configurations. Gas phase (——); aqueous solution (-----). Values were computed at the PBE0/[6-311++G(2d,2p), 6-31+G(d,p)] level. Aqueous solution data were computed including the two closest water molecules in the QM layer.

Table 8. Isotropic shielding constants [ppm] of acetone in the gas phase and in aqueous solution. PBE0 and MP2 calculations were carried out with the [6-311++G(2d,2p), 6-31+G(d,p)] basis set. Aqueous solution data were computed including the two closest water molecules in the QM layer. Gas-phase and aqueous solution results are averaged over 100 configurations.

		PBE0	MP2	Experiment
^{13}C	gas-phase	−23	−9	–
	solution	−42	−28	–
	$\Delta_{solvent}$	−19	−19	−18.9[42]
^{17}O	gas-phase	−340	−265	
	solution	−253	−200	
	$\Delta_{solvent}$	87	65	75.5[43]

Finally, there is satisfactory agreement between the computed and experimental results for the $n\rightarrow\pi^*$ excitation energy in the gas phase (4.38 eV from our computations and 4.5 eV from ref. [41]) and, especially, its solvent shift (0.19 eV from our computations and 0.2 eV from ref. [41]).

4. Conclusions

Herein, a reliable and computationally efficient approach has been proposed for the theoretical evaluation of spectroscopic observables of molecules in solution. Such a challenging problem requires that solvent effects and thermal fluctuations of the molecular conformations should be properly taken into account. Our procedure has been schematically divided into three successive steps: 1) the parameterization of an effective classical force field for the solute intramolecular and solute/solvent intermolecular interactions, 2) the sampling of the configurational space of the system via MD simulations using the mean field (MF) model,[24] and 3) the QM computation of spectroscopic parameters on a representative set of molecular configurations. This computational approach has been illustrated by considering the acetone molecule in aqueous solution as a test system, which has been extensively studied both experimentally[41–43] and, more recently, theoretically[16,18,27,28,43] by different groups.

First, we have shown that classical MD based on a reliable force field compares favorably with CPMD simulations, as shown by acetone/water RDFs and hydrogen-bond distributions, and can be used effectively to obtain extended sampling of the solute/solvent configurations. To remark once more on the importance of point 1, it should be noted that it was necessary to modify the standard OPLS-AA force field[21] for acetone by including oxygen lone pairs and to improve some geometrical features, such as the C=O stretching and out-of-plane mode, by an accurate reparameterization. Indeed, force-field-based MD could provide more reliable molecular geometries, and indirectly spectroscopic properties as seen for ^{13}C and ^{17}O shielding constants, with respect to standard Car–Parrinello MD, due to the well-known limitations of the DFT functionals used in the latter. Besides, the MF model has been shown to be an effective methodology for simulating solute/solvent systems, and is particularly well-suited to post ab initio calculations of spectroscopic parameters.

Then, it was shown how a mixed QM/MM approach can accurately reproduce experimental spectroscopic observables of acetone, such as NMR shielding constants and $n\rightarrow\pi^*$ excitation energy, provided that an appropriate level of theory and a suitable number of solvent molecules are treated at the QM level along with the solute. In particular, the present approach has the advantage of being highly flexible concerning the QM model used for the post-MD calculations, that is, DFT or post-Hartree–Fock methods, and was found to be equivalent to more expensive full QM calculations as well as to QM calculations using the PCM model, as shown for the case of acetone ^{17}O NMR chemical shifts. Moreover, due to a much longer simulation timescale, our procedure has provided results with a significantly smaller statistical error for the computed UV transition energies than that for previous results based on full QM[18,28] and QM/MM[16] molecular dynamics simulations.

Currently, we are planning to apply the proposed computational strategy to study systematically the spectroscopic properties of a variety of molecular systems in solution, from small to large molecules.

Acknowledgements

The authors thank the Italian Ministry for University and Research (MIUR) and Gaussian Inc. for financial support, and the Campus Grid at Naples for computer facilities.

Keywords: ab initio calculations · acetone · molecular dynamics · NMR spectroscopy · UV/Vis spectroscopy

[1] a) T. Helgaker, M. Jaszunski, K. Ruud, *Chem. Rev.* **1999**, *99*, 293; b) M. Merchan, L. Serrano-Andres, M. P. Fulscher, B. O. Roos, *Recent Adv. Comput. Chem.* **1999**, *4*, 161; c) B. O. Roos, *Acc. Chem. Res.* **1999**, *32*, 137; d) R. Improta, V. Barone, *Chem. Rev.* **2004**, *104*, 1231.

[2] a) S. Goedecker, G. E. Scuseria, *Comput. Sci. Eng.* **2003**, *5*, 14; b) G. E. Scuseria, *J. Phys. Chem. A* **1999**, *103*, 4782.

[3] a) P. Hohenberg, W. Kohn, *Phys. Rev.* **1964**, *136*, B864; b) W. Kohn, L. J. Sham, *Phys. Rev.* **1965**, *140*, A1133.

[4] a) E. K. U. Gross, W. Kohn, *Phys. Rev. Lett.* **1985**, *55*, 2850; b) R. E. Stratmann, G. E. Scuseria, M. J. Frish, *J. Chem. Phys.* **1998**, *109*, 8218; c) M. E. Casida in *Recent Advances in Density Functional Methods, Part I* (Ed.: D. P. Chong), World Scientific, Singapore, **1994**, pp. 155–192.

[5] a) K. Wolinski, J. F. Hilton, P. Pulay, *J. Am. Chem. Soc.* **1990**, *112*, 8251; b) J. R. Cheeseman, G. W. Trucks, T. A. Keith, M. J. Frish, *J. Chem. Phys.* **1998**, *104*, 5497.

[6] O. Crescenzi, G. Correale, A. Bolognese, V. Piscopo, M. Parrilli, V. Barone, *Org. Biomol. Chem.* **2004**, *1*, 1577.

[7] a) A. Bagno, F. Rastrelli, G. Sacelli, *J. Phys. Chem. A* **2003**, *107*, 9964; b) D. J. Giesen, N. Zumbulyadis, *Phys. Chem. Chem. Phys.* **2002**, *4*, 5498; c) C. Benzi, M. Cossi, V. Barone, *Phys. Chem. Chem. Phys.* **2004**, *6*, 2557; d) C. Benzi, O. Crescenzi, M. Pavone, V. Barone, *Magn. Reson. Chem.* **2004**, *42*, S57.

[8] a) M. Persico, J. Tomasi, *Chem. Rev.* **1994**, *94*, 2027; b) C. J. Cramer, D. G. Truhlar, *Chem. Rev.* **1999**, *99*, 2161; c) J. Tomasi, B. Mennucci, R. Cammi, *Chem. Rev.* **2005**, *105*, 2999.

[9] a) S. Iwata, K. Morokuma, *J. Am. Chem. Soc.* **1975**, *97*, 966; b) M. Kraus, S. P. Webb, *J. Chem. Phys.* **1997**, *107*, 5771.

[10] J. Gauss, *J. Chem. Phys.* **1993**, *99*, 3629; b) J. Gauss, *Phys. Chem. Chem. Phys.* **1995**, *99*, 1001; c) R. E. Stratmann, G. E. Scuseria, M. J. Frisch, *J. Chem. Phys.* **1998**, *109*, 8218.

[11] S. Miertus, E. Scrocco, J. Tomasi, *Chem. Phys.* **1981**, *55*, 117.

[12] a) M. Cossi, G. Scalmani, N. Rega, V. Barone, *J. Chem. Phys.* **2002**, *117*, 43; b) G. Scalmani, V. Barone, K. N. Kudin, C. S. Pomelli, G. E. Scuseria, M. J. Frisch, *Theor. Chem. Acc.* **2004**, *111*, 90; c) V. Barone, M. Cossi, J. Tomasi, *J. Chem. Phys.* **1997**, *107*, 3210; d) M. Cossi, V. Barone, *J. Phys. Chem. A* **2000**, *104*, 10614.

[13] a) C. Adamo, M. Cossi, N. Rega, V. Barone in *Theoretical Biochemistry: Processes and Properties of Biological Systems, Theoretical and Computational Chemistry, vol. 9*, Elsevier, New York, **1999**; b) M. Cossi, V. Barone, *J. Chem. Phys.* **2000**, *112*, 2427; c) M. Cossi, V. Barone, *J. Chem. Phys.* **2001**, *115*, 4708.

[14] a) C. Adamo, V. Barone, *Chem. Phys. Lett.* **2000**, *320*, 152; b) F. Aquilante, B. Roos, V. Barone, *J. Chem. Phys.* **2003**, *119*, 12323.

[15] a) S. Canuto, K. Coutinho, M. Zerner, *J. Chem. Phys.* **2000**, *112*, 7293; b) K. Coutinho, S. Canuto, *J. Chem. Phys.* **2000**, *113*, 9132; c) K. Coutinho, S. Canuto, *Theochem* **2003**, *632*, 235.

[16] a) U. F. Rohrig, I. Frank, J. Hutter, A. Laio, J. VandeVondele, U. Rothlisberger, *ChemPhysChem* **2003**, *4*, 1177; b) M. Sulpizi, P. Carloni, J. Hutter, U. Rothlisberger, *Phys. Chem. Chem. Phys.* **2003**, *5*, 4798; c) M. Sulpizi,

U. F. Rohrig, J. Hutter, U. Rothlisberger, *Int. J. Quantum Chem.* **2005**, *101*, 671.
[17] R. Car, M. Parrinello, *Phys. Rev. Lett.* **1985**, *55*, 2471.
[18] L. Bernasconi, M. Sprik, J. Hutter, *J. Chem. Phys.* **2003**, *119*, 12417.
[19] a) C. Adamo, A. di Matteo, V. Barone, *Adv. Quantum Chem.* **1999**, *36*, 45; b) M. Pavone, V. Barone, I. Ciofini, C. Adamo, *J. Chem. Phys.* **2004**, *120*, 9167.
[20] D. J. Searles, H. Huber, *Encyclopedia of Nuclear Magnetic Resonance, vol. 9: Advances in NMR*, pp. 215–226, John Wiley & Sons, Chichester, **2002**; b) D. J. Searles, H. Huber in *Calculation of NMR and EPR Parameters: Theory and Applications* (Eds.: M. Kaupp, M. Bühl, V. G. Malkin), Wiley-VCH, Weinheim, **2004**, pp. 175 – 189.
[21] D. Kony, W. Damm, S. Stoll, W. F. Van Gunsteren, *J. Comput. Chem.* **2002**, *23*, 1416.
[22] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, Jr., D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, P. A. Kollman, *J. Am. Chem. Soc.* **1995**, *117*, 5179.
[23] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, M. Karplus, *J. Comput. Chem.* **1983**, *4*, 187.
[24] G. Brancato, A. Di Nola, V. Barone, A. Amadei, *J. Chem. Phys.* **2005**, *122*, 154109.
[25] a) D. Sebastiani, M. Parrinello, *J. Phys. Chem. A* **2001**, *105*, 1951; b) M. Bühl, *J. Chem. Phys. A* **2002**, *106*, 10505.
[26] a) M. Aschi, R. Spezia, A. Di Nola, A. Amadei, *Chem. Phys. Lett.* **2001**, *344*, 374; b) R. Spezia, M. Aschi, A. Di Nola, A. Amadei, *Chem. Phys. Lett.* **2002**, *365*, 450; c) A. Amadei, F. Marinelli, M. D'Abramo, M. D'Alessandro, M. Anselmi, A. Di Nola, M. Aschi, *J. Chem. Phys.* **2005**, *122*, 124506.
[27] K. Aidas, J. Kongsted, A. Osted, K. V. Mikkelsen, O. Christiansen, *J. Phys. Chem. A*, **2005**, *109*, 8001.
[28] O. Crescenzi, M. Pavone, F. De Angelis, V. Barone, *J. Phys. Chem. B* **2005**, *109*, 445.
[29] J. P. Perdew, K. Burke, M. Ernzerhof, *Phys. Rev. Lett.* **1996**, *77*, 3865.
[30] C. Adamo, V. Barone, *J. Chem. Phys.* **1999**, *110*, 6158.
[31] M. J. Frisch et al., Gaussian 03, revision C.02, Gaussian, Inc., Pittsburgh, PA, **2004**.
[32] P. Cieplak, W. D. Cornell, C. Bayly, P. A. Kollman, *J. Comput. Chem.* **1995**, *16*, 1357.
[33] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gusteren, J. Hermans in *Intermolecular Forces* (Ed.: B. Pullman), Reidel, Dordrecht, **1981**.
[34] D. V. der Spoel et al., Gromacs User Manual version 3, Nijenborgh 4, 9747 AG Groningen, The Netherlands, **2001**.
[35] A. Amadei, G. Chillemi, M. A. Ceruso, A. Grottesi, A. Di Nola, *J. Chem. Phys.* **2000**, *112*, 9.
[36] M. P. Allen, D. J. Tildesly, *Computer Simulation of Liquids*, Oxford University Press, Oxford, **1987**.
[37] M. M. Francl, W. J. Petro, W. J. S. Hehre, J. Binkley, M. S. Gordon, D. J. DeFrees, J. A. Pople, *J. Chem. Phys.* **1982**, *77*, 3654.
[38] C. Moller, M. S. Plesset, *Phys. Rev.* **1934**, *46*, 618.
[39] R. L. Hilderbrandt, A. L. Andreassen, S. H. Bauer, *J. Phys. Chem.* **1970**, *74*, 1586.
[40] T. H. Dunning, Jr., *J. Chem. Phys.* **1989**, *90*, 1007.
[41] a) N. S. Bayliss, E. G. McRae, *J. Phys. Chem.* **1954**, *58*, 1006; b) A. Balasubramanian, C. N. R. Rao, *Spectrochim. Acta* **1962**, *18*, 1337; c) N. S. Bayliss, G. Wills-Johnson, *Spectrochim. Acta A* **1968**, *24*, 551.
[42] B. Tiffon, J. E. Dubois, *Org. Magn. Reson.* **1978**, *11*, 295.
[43] M. Cossi, O. Crescenzi, *J. Chem. Phys.* **2003**, *118*, 8863.

Received: July 6, 2005
Published online on December 6, 2005

