FEATURE ARTICLE

Ab Initio Molecular Dynamics Computation of the Infrared Spectrum of Aqueous Uracil

Marie-Pierre Gaigeot[†]

Laboratoire de Physicochimie Biomoléculaire et Cellulaire, UMR CNRS 7033, Université Pierre et Marie Curie, 4 Place Jussieu, Case courrier 138, F-75252 Paris Cedex 05, France

Michiel Sprik*

Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, United Kingdom

Received: March 27, 2003; In Final Form: July 16, 2003

Recent progress in the development of ab initio molecular dynamics methods for the computation of infrared absorption spectra in condensed molecular systems is reviewed and illustrated by a detailed account of an application to aqueous uracil. Similar to classical force field simulations, the spectrum is obtained as the Fourier transform of the polarization time autocorrelation function. The density functional methodology for the computation of electronic polarization in periodic supercells is briefly outlined, and also the effect of quantum corrections is discussed. The spectral patterns obtained for the model system in the 2000–1000 cm^{-1} domain are in good agreement with experiment. Comparing to the low-temperature vacuum spectrum computed by similar time-dependent methods, we found that the narrow amide bending band in a vacuum is spread out over a 500 cm^{-1} wide interval in solution with a substantially blue-shifted high-frequency end. The highest increase in frequency was found for N1–H1 bending. The red shift and broadening of C=O stretching bands is, in comparison, a much smaller effect.

1. Introduction

Ab Initio Molecular Dynamics of Aqueous Systems. The quantum chemistry and statistical mechanics of molecular liquids are two branches of theoretical chemistry that are of vital importance in physical chemistry. Methodology and computational demands of these two disciplines are, in practice, rather different. As a result, they are usually separated by an intermediate modeling step transforming the results of the electronic structure calculation in parametrized model potentials. In the ab initio molecular dynamics (MD) method,[1] the nuclear gradients obtained from a continuously updated electronic structure computation are directly applied as mechanical forces. Construction and parametrization of models is avoided. Electronic structure calculation and finite temperature statistical mechanics are unified at a fundamental level in ab initio MD.

First-principle rigor, of course, comes at a considerable computational cost, limiting the size of model systems and the time span over which they can be followed. Applications in physical chemistry for which the extra expense of direct dynamics has paid off typically involve ionic or polar liquids characterized by strong coupling between electronic structure and thermal fluctuations. A prominent example is the chemistry of aqueous ionic solutions. Classical reactions such as acid–base and redox equilibria of simple ions fall in this category

but also a large fraction of the complex reactions of biochemistry. To date, ab initio MD studies of aqueous reactions have employed almost exclusively the Car–Parrinello method[1] (for an extensive technical introduction see ref 2). The electronic structure component in this approach is based on the plane wave–pseudopotential implementation of the density functional theory (DFT) of extended systems. This method, originally developed in solid-state physics, is combined with extended Lagrangian molecular dynamics techniques borrowed from classical liquid-state simulation.[3] With the computational resources commonly available to academic research groups, the system dimensions and time scales accessible using this methodology are 10–20 ps trajectories of model systems consisting of 50 solvent molecules and one or more small solute molecules.

A time window of 10–20 ps is just long enough to observe a number of important dynamical processes characteristic for liquid water, in particular hydrogen-bond breaking and rearrangement (∼2–3 ps) and hopping of excess protons in acids (∼1–2 ps). Moreover, in model systems of 30–50 H_2O molecules (corresponding to cubic periodic cells of length 9–12 Å), small solute molecules can be surrounded by a solvation shell sufficiently large to adopt all of the often specific structures relevant to the aqueous chemistry of the solute. Thus, after it had been established that the structural and dynamical properties of liquid water can be reproduced by applying the ab initio MD method to these minimal models (provided a suitable density functional is used),[4,5] a number of ab initio MD studies of elementary aqueous chemical reactions followed, such as the

* To whom correspondence should be addressed. E-mail: ms284@cam.ac.uk.
† Present address: Laboratoire de Modelisation de Systemes Moleculaires Complexes, Université d' Evry val d'Essone, Rue de Pere A. Jarland, 91405 Evry cedex, France. E-mail: gaigeot@ccr.jussieu.fr.

Feature Article

J. Phys. Chem. B, Vol. 107, No. 38, 2003 **10345**

structure and transport of excess protons (hydronium ion)[6,7] and missing protons (hydroxyl ion),[8] water autodissociation,[9-11] acid dissociation,[12,13] and other model reactions.[14-16] In parallel, it was verified that similar techniques are also capable of describing the hydration of polar molecules, such as methanol[17] and DMSO,[18] of aqua ions, such as alkali,[19-22] alkaline earth,[23,24] and (post)transition metal cations,[21-26] ammonium[27] and aluminum hydroxide,[28] and halide[12,13,29] and oxyanions.[30-32] These studies have prepared the way for the application of Car−Parrinello techniques to larger molecules of biological interest immersed in liquid water.[33-37] Car−Parrinello simulation of biological systems has since developed into a very active field and is, arguably, one of the most promising applications of this method. For an impression of the state of the art, we refer to the review papers by Carloni et al. (see, for example, ref 38).

First-Principle Infrared Spectroscopy in Solution. The topic of the present contribution is another application of Car−Parrinello simulation crucially dependent on coupling between electronic structure and molecular dynamics, namely, the computation of molecular vibrational spectra of finite temperature condensed systems. In classical modeling, spectroscopic calculations are well established and have already a long history. Molecular dynamics simulation, combined with time correlation function methods, has made vital contributions to the understanding of the optical response of liquids and solutions (see, for example, refs 39−48). In the ab initio context, this combination of computational techniques is of much more recent date. The pioneering study in the field is a first-principle calculation of the infrared (IR) absorption of liquid water[49] and ice at high pressure[50] by Silvestrelli, Bernasconi, and Parrinello, followed by the computation of the Raman spectrum of ice by Putrino and Parrinello[51] using density functional perturbation methods.[52]

This development has been made possible by the major advances in the electronic structure calculation of extended periodic systems due to Vanderbilt and King-Smith.[53,54] Their key contribution is a derivation of an expression for the electronic polarization of insulating crystalline solids in terms of a geometric (Berry) phase associated with the ground-state wave function thus avoiding the paradox created by electron density based pictures of dipole moments in periodic systems.[55] An alternative formulation of the connection between polarization and a Berry phase, starting from a definition of position in periodic systems, was given by Resta.[56-58] This approach is most convenient for the computation of polarization in supercells of disordered systems and was applied by Silvestrelli et al.[49] in their study of the IR absorption of liquid water. Similar to previous classical computational investigations, the spectrum was obtained by Fourier transformation of the time correlation function of total dipole moment. However, unlike in classical model studies, no separation between permanent and induced molecular dipole moments is needed in ab initio MD. All of these effects are automatically contained in the total electronic polarization.

The spectroscopy of aqueous solutions adds further challenges. Because of the intense IR activity of water, the signal of the hydrated solute is usually completely obscured. This is a problem in experiment, as well as computation (see, for example, the study of the IR spectrum of aqueous *N*-methylacetamide (NMA) of ref 48). A theorist, of course, has a variety of technical options for filtering out solvent contributions not available to experimentalists. Exploiting this freedom, we will make use of a modification of the Vanderbilt method, due to Marzari and Vanderbilt.[59] Building on the new insights in

polarization of crystals, Marzari and Vanderbilt developed a procedure for optimizing the degree of localization of Wannier functions. They also showed that the total polarization can be separated in a sum of contributions from the centers of these maximally localized Wannier orbitals. In condensed molecular systems, the overlap between Wannier orbitals localized on different molecules is small because intermolecular interactions are nonbonding. The maximally localized Wannier function scheme can therefore be used to separate the total charge density of condensed molecular systems in molecular contributions. Silvestrelli and Parrinello[5,60] used this idea for an estimation of the effective dipole moment of a water molecule in the liquid. The Wannier function description was also the basis for the analysis of the response of electronic polarization and infrared absorption of high-pressure hydrogen of ref 61. We will adopt the same approach to isolate the infrared signal of aqueous organic solutes.

Application to Aqueous Uracil. Nucleic acids are exposed to water solvent in a number of biological structures. The hydration of nucleic acids is therefore of great interest and has been investigated extensively by infrared spectroscopy.[62-67] Because of their role as building blocks, also single bases solvated in water have been the subject of spectroscopic study.[68-70] These relatively small aqueous molecules are very suitable targets for ab initio molecular dynamics simulation methods, and therefore, we chose the computation of the infrared spectrum of uracil, the smallest of nucleic bases, to illustrate the potential of "first-principle" vibrational spectroscopy as a new tool in the computational chemistry of complex molecular systems.

The aqueous uracil model system used in this work is shown in Figure 1. It consists of one solute molecule surrounded by 49 water molecules (total number of 159 atoms). The dimensions of the periodically repeated cubic cell are 11.5 Å in all three directions. With a solvent system of this size, two complete layers of solvent and part of the third hydration shell around the solute are taken into account. The dynamics of this system was followed over a period of 7.3 ps. The first piece of information that can be obtained from a molecular dynamics trajectory of this length is a fairly accurate description of the structural and dynamical properties of the solvent surrounding the uracil molecule and the effect on the solute geometry. These results will be reported in a separate publication.[71] The present paper focuses on the computation of electric dipole moments, vibrational motion, and infrared spectrum using the same microscopic data. The positions, intensities, and band shapes are compared in detail to experiment[70] and interpreted with the help of the insights in the hydration patterns obtained from the structural analysis of ref 71.

Organization of this paper is as follows. We start with an outline of the computational methodology in section 2. In section 3, we discuss the application to uracil in vacuo. This calculation serves as a benchmark and validation of the ab initio molecular dynamics approach. Then after a summary of the structural and hydration properties of aqueous uracil in section 4 (discussed in more detail in ref 71), we present our results on the static polarization (section 5) and vibrations and the infrared spectrum (section 6).

2. Computational Methods

Polarization and Infrared Absorption in Periodic Systems. The formalism of dipole−dipole time correlation functions for extracting infrared spectra from molecular dynamics simulations can be found in most text books on statistical mechanics (see,
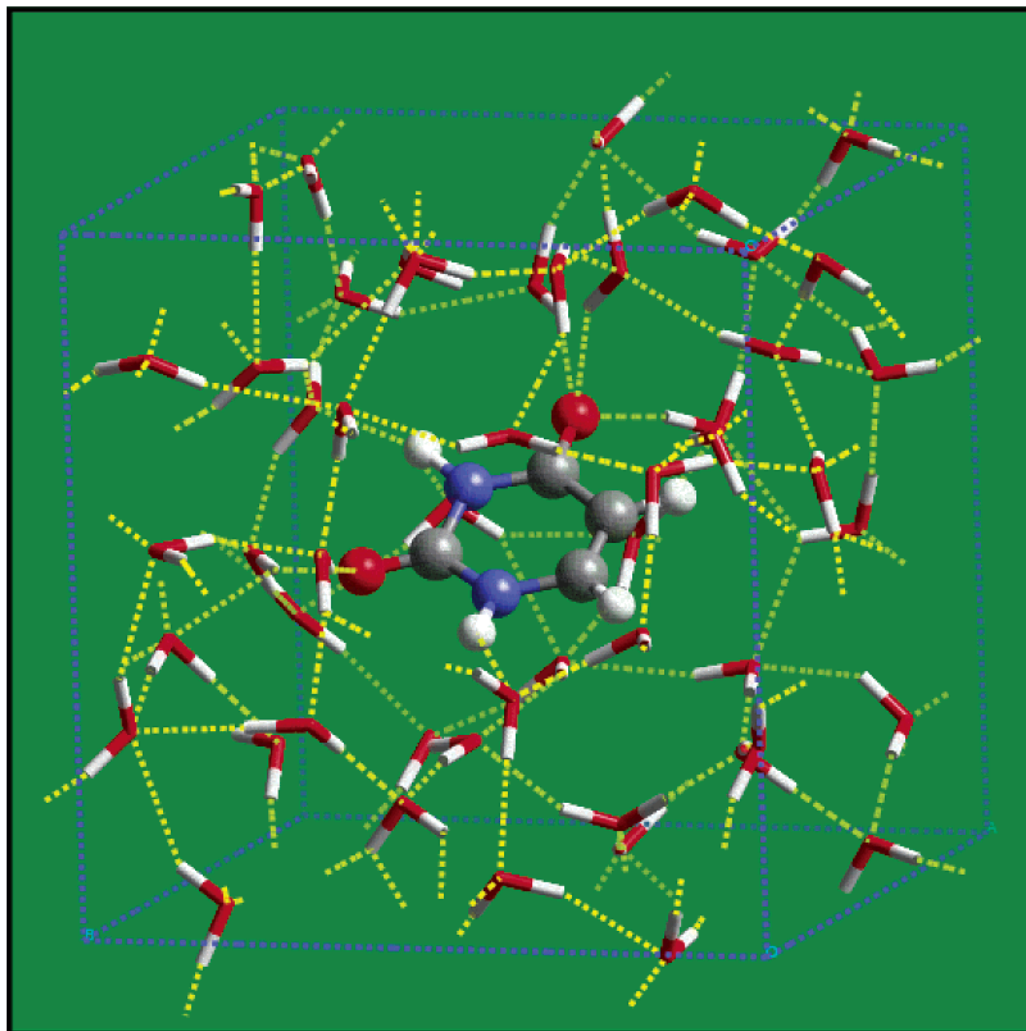
Figure 1. Instantaneous configuration of uracil surrounded by 49 water molecules sampled from the ab initio molecular dynamics simulation. The cubic box size is 11.5 Å. Periodic boundary conditions are applied.

for example, ref 72). The infrared absorption coefficient, $\alpha(\omega)$, is calculated from the Fourier transform of the time correlation of the total dipole moment, $\mathbf{M}(t)$, of the system according to

$$\alpha(\omega)n(\omega) = \frac{2\pi\omega(1 - \exp(-\beta\hbar\omega))D(\omega)}{3\hbar cV} \times$$
$$\int_{-\infty}^{\infty} dt \, \langle \mathbf{M}(t)\cdot\mathbf{M}(0)\rangle \exp(i\omega t) \quad (1)$$

where $\beta = 1/(kT)$, T being the temperature and k the Boltzmann constant. c is the speed of light in a vacuum, V is the volume, and $n(\omega)$ is the refractive index of the medium at frequency ω. Brackets indicate equilibrium ensemble averages, which in simulation are evaluated as a time average over the length of the dynamics. $D(\omega)$ in eq 1 is a quantum correction factor, multiplying the classical line shape to correct for the violation of the detailed balance conditions by the classical treatment and to account for zero-point motion effects related to the presence of hydrogen atoms. There is an unfortunate degree of arbitrariness in the choice for $D(\omega)$. The functional forms of the various options and their effect on the intensities will be discussed in the Appendix.

When calculating infrared absorption in the gas phase, $\mathbf{M}(t)$ in eq 1 is simply the dipole moment of a molecule. In case of liquids and solids, $\mathbf{M}(t)$ is the total dipole moment of the supercell modeling the extended system. In classical simulation

of molecular liquids, $\mathbf{M}(t)$ can be computed by adding up the dipoles of the molecules as defined by the force field model (e.g., see ref 41). In electronic structure calculation, such an approach is only possible when the overlap between the molecular charge distributions is negligible (Clausius−Mossotti limit), a condition which is strictly valid only in the gas phase. Partitioning schemes of charge density in the condensed phase are frustrated by the periodic boundary conditions. In fact, as underlined in ref 55, such schemes are "in principle" inconsistent because the polarization of a bulk sample is determined by the charge distributions at the (physical) boundaries, which have been eliminated in periodic models. This paradox has been solved in the modern theory of polarization.[53,54] Focusing on the transient currents associated with changes in polarization, Vanderbilt and King-Smith were able to relate this property to the phase of the ground-state wave function rather than the charge density. This led them to their famous Berry phase expression for polarization.

The Berry phase expression of refs 53 and 54 was derived in terms of integrals and derivatives of Bloch functions in k space. However, k-space sampling in ab initio MD simulation of disordered systems is usually restricted to $\mathbf{k} = 0$ (the Γ point) and a single k-point approximation must be applied. Resta's real space formulation of the Berry phase representation[58] seems more suitable for this purpose, and the expression for polarization used by Silvestrelli et al. in their study of the IR absorption

Feature Article

J. Phys. Chem. B, Vol. 107, No. 38, 2003 **10347**

of liquid water[49] is based on this approach. The central quantity in the Resta picture is the dimensionless complex number (see also ref 73)

$$z_N = \langle \Psi | e^{-i\mathbf{G}_\alpha \cdot \hat{\mathbf{R}}} | \Psi \rangle \qquad (2)$$

where $\mathbf{G}_\alpha$, $\alpha = 1, 2, 3$, is a reciprocal lattice vector of the simple cubic supercell of length L ($\mathbf{G}_1 = 2\pi/L(1,0,0)$, $\mathbf{G}_2 = 2\pi/L(0,1,0)$, $\mathbf{G}_3 = 2\pi/L(0,0,1)$) and $\hat{\mathbf{R}} = \sum_{i=1}^N \hat{\mathbf{r}}_i$ denotes the collective position operator of the N electrons. Ψ is the ground-state wave function (either the fully correlated wave function or a Slater determinant in an independent electron approximation). Resta has shown that in the limit where the Γ point approximation applies (large L) the electronic contribution to the cell dipole moment, M_α^{el}, is given by

$$M_\alpha^{\text{el}} = \frac{e}{|\mathbf{G}_\alpha|} \text{ Im ln } z_N \qquad (3)$$

The quantity Im ln z_N is the Berry phase, which in terms of a set of occupied Kohn–Sham orbitals, $\psi_k(\mathbf{r})$, is computed according to

$$\text{Im ln } z_N = 2 \text{ Im ln det } \mathbf{S} \qquad (4)$$

with the matrix $\mathbf{S}$ given by

$$S_{kl} = \langle \psi_k | e^{-i\mathbf{G}_\alpha \cdot \hat{\mathbf{r}}} | \psi_l \rangle \qquad (5)$$

where we have assumed a spin-restricted description.

In the present study of infrared spectra of aqueous solutions, we adopt an alternative route that makes use of the maximally localized Wannier function scheme.[59] We were inspired by the computation of the effective molecular dipole moment of water molecules in the liquid by Silvestrelli and Parrinello.[5,60] A similar approach was applied in ref 61 in a study of molecular hydrogen under high pressure. Building on the breakthrough in the treatment of polarization in periodic systems, Marzari and Vanderbilt developed a method for constructing a set of optimally localized Wannier functions from the Bloch states in a crystalline solid.[59,74] Wannier functions, $w_n(\mathbf{r})$, are formally related to the Kohn–Sham orbitals, $\psi_k(\mathbf{r})$, by a unitary transformation

$$w_n(\mathbf{r}) = \sum_{k=1}^{N/2} U_{nk} \psi_k(\mathbf{r}) \qquad (6)$$

and are therefore orthogonal (as in eq 4, we have assumed that we are dealing with a system of N perfectly paired electrons described by $N/2$ spin-restricted one-electron orbitals). The properties of the Wannier functions, such as their location and density (square) overlap, depend on the choice of the transformation $\mathbf{U}$, which is to a large extent arbitrary. Marzari and Vanderbilt removed this arbitrariness by imposing restrictions on the degree of localization. The key step was the generalization of the new understanding of polarization to a definition of position in periodic systems, which could be used to specify the spatial width of Wannier states. Localization can then be optimized by finding the unitary matrix $\mathbf{U}$ in eq 6 minimizing the total spread

$$\Omega = \sum_{n=1}^{N/2} (\langle w_n | r^2 | w_n \rangle - \langle w_n | \mathbf{r} | w_n \rangle^2) \qquad (7)$$

The maximally localized Wannier states obtained by this procedure can thus be compared to the Boys localized orbitals

of quantum chemistry.[75] The expectation value of position $\mathbf{r}_n = \langle w_n | \mathbf{r} | w_n \rangle$ in eq 7 can be interpreted as the location of the center of the Wannier function w_n and is often referred to as a Wannier function center (WFC). Similar to polarization, Wannier functions in disordered systems can be treated using the real space approach of Resta. The first calculation of this kind was carried out by Silvestrelli and co-workers.[76] The expression for the x component of a WFC closely resembles eq 3 for the dipole moment in the x direction

$$x_n = -\frac{L}{2\pi} \text{ Im ln} \langle w_n | e^{-i2\pi x/L} | w_n \rangle \qquad (8)$$

The corresponding expression for the average of the squared position vector under periodic boundary conditions can be found in ref 73. In applications to large supercells,[76] however, evaluation of this quantity is avoided. Provided the cell is large compared to the width, the same set of maximally localized Wannier functions can be obtained by maximizing the function

$$\tilde{\Omega} = \sum_{n=1}^{N/2} (|X_{nm}|^2 + |Y_{nm}|^2 + |Z_{nm}|^2) \qquad (9)$$

with $X_{nm} = \langle w_n | e^{-i2\pi x/L} | w_m \rangle$ (similar definitions for Y_{nm} and Z_{nm} apply,[76] see also refs 77 and 78). This is how the construction of maximally localized Wannier functions is implemented in the CPMD package used in the present work (see below).

The electronic contribution to the polarization (or rather the cell dipole moment) can be directly computed from the position of the WFC of eq 8.

$$\mathbf{M}^{\text{el}} = -2e \sum_{n=1}^{N/2} \mathbf{r}_n \qquad (10)$$

with $\mathbf{r}_n = (x_n, y_n, z_n)$ denoting the position vector of the WFC corresponding to w_n. As demonstrated in ref 59, the expression in terms of WFCs is identical to the polarization obtained from the Berry phase when evaluated using full k space integration. In fact, this identity holds even for general Wannier functions, which are not maximally localized. This is no longer true in the single-point scheme. However, under the condition that the width of the Wannier functions is small compared to the size L of the supercell, eq 10 remains a good approximation, which is why the use of the WFCs of maximally localized Wannier functions is mandatory for the computation of polarization of disordered systems.[76]

Solute Spectrum and Vibrational Assignment. The parallel of expression of eq 10 to the familiar classical picture of polarization as a sum over coordinate vectors of point charges is very suggestive. Indeed the maximally localized Wannier functions of a system of closed-shell molecules tend to be located on or near bonds and atomic centers. As a result, it is possible to associate with each molecule as many WFCs as the number of valence orbitals. This can be done on the basis of a simple distance criterion. A similar approach has already been used in a number of ab initio MD studies of molecular liquids and solids[5,17,24,60,61] and proven to be a most powerful tool for recovering a molecular picture from electronic structure calculations. Here the same approach will be applied to separate the total dipole moment in solute and solvent contributions. If the dipole of the solute (uracil) is denoted by $\boldsymbol{\mu}_U$ and the dipole moment of water molecule I by $\boldsymbol{\mu}_I$, the dipole time correlation

function (eq 1) can be decomposed into three parts:

$$C_{MM}(t) = C_{WW}(t) + C_{UU}(t) + C_{UW}(t) \qquad (11)$$

where

$$C_{WW}(t) = \sum_{I,J}^{solvent} \langle \boldsymbol{\mu}_I(t) \cdot \boldsymbol{\mu}_J(0) \rangle \qquad (12)$$

$$C_{UU}(t) = \langle \boldsymbol{\mu}_U(t) \cdot \boldsymbol{\mu}_U(0) \rangle \qquad (13)$$

$$C_{UW}(t) = \sum_{I}^{solvent} \langle \boldsymbol{\mu}_I(t) \cdot \boldsymbol{\mu}_U(0) \rangle + \sum_{I}^{solvent} \langle \boldsymbol{\mu}_U(t) \cdot \boldsymbol{\mu}_I(0) \rangle \quad (14)$$

When Fourier transformed, the first term (eq 12) gives the infrared spectrum of the water solvent, while the second and third terms (eqs 13 and 14) contribute to the infrared spectrum of uracil.

While intuitively appealing, separation of polarization into solvent and solute contributions remains problematic for reasons that have become abundantly clear in the light of the development of the modern theory of polarization. Indeed, the most direct way to study infrared absorption of aqueous molecules is subtracting from the full solution spectrum a suitably normalized background determined from a pure solvent calculation. This would also be consistent with the way most experimental data have been obtained. However, this scheme would confront us with serious signal-to-noise problems (even worse than those in experiment because of the smaller samples and shorter time span of simulations). For this reason (and for analysis purposes), we will focus on the solute dipole autocorrelation of eq 13 as a probe of the solvent effect on the infrared spectrum of uracil. The information contained in the cross-correlation term (eq 14) will not be discussed. Unlike the autocorrelation spectrum, the contribution due to cross-correlations can be both positive and negative, making a determination of comparable accuracy more difficult. Moreover this term will be also more affected by the ambiguities in our definition of molecular dipoles, which as explained above ultimately has no firm basis in theory or experiment. To address these important issues in a more rigorous way, it is probably better to abandon the dipole representation of electromagnetic coupling and return to the more fundamental current representation.

The area of molecular spectroscopy in which computation is perhaps most useful is interpretation and assignment. For medium-sized molecules in a vacuum, a full normal-mode analysis of the vibrational dynamics is available, making decomposition of the spectrum relatively straightforward. Such an analysis is not feasible for finite temperature liquids, and we must resort to dynamical correlation methods. The key quantity here is the vibrational density of states (VDOS) as obtained by Fourier transformation of time correlation functions of positions and velocities. Assignments of the vibrational peaks in the VDOS in terms of atomic internal displacements were carried out by matching the peaks in the VDOS spectrum to selected $\langle \mathbf{a}(0) \cdot \mathbf{b}(t) \rangle$ bond vector time correlation functions. The vector $\mathbf{a}(0)$ at the reference time $t = 0$ is calculated in the optimized geometry, whereas $\mathbf{b}(t)$ is sampled from the configuration generated at time t of the dynamics. We found that stretching motions could be traced best using $\langle \mathbf{a}(0) \cdot \mathbf{a}(t) \rangle$ autocorrelation functions while for bending motions $\langle \mathbf{a}(0) \cdot \mathbf{b}(t) \rangle$ cross-correlation functions were more convenient. For instance, the amide N−H stretching motion was assigned by comparing the total VDOS to the power spectrum of $\langle \mathbf{R}_{NH}(0) \cdot \mathbf{R}_{NH}(t) \rangle$,

where $\mathbf{R}_{NH}$ is the vector connecting a pair of bonded N and H atoms in uracil. The amide N−H bending motion was identified by searching for features in common with the power spectrum of $\langle \mathbf{d}_{NH}(0) \cdot \mathbf{R}_{NH}(t) \rangle$, where $\mathbf{d}_{NH}(0)$ is the vector perpendicular to $\mathbf{R}_{NH}(0)$ contained in the molecular plane in the optimized geometry of uracil. This admittedly crude method allows us to see how the different local vibrational modes are distributed over the total vibrational density of states. It cannot quantify the relative percentages of participation of the different internal motions to each vibrational mode. However, by comparing the bands appearing in the vibrational density of states to the actual infrared spectrum, we are still able to determine which vibrational modes are infrared active, that is, sensitive to variations of the dipole moment of the molecule.

Ab Initio Molecular Dynamics. The technical implementation of the density functional calculation is the same as that in previous ab initio MD studies of aqueous systems. After numerous validations, this setup has become standard for the application of the Car−Parrinello method to first-row chemistry. The one-electron orbitals are expanded in a plane wave basis set with a kinetic energy cutoff of 70 Ry restricted to the Γ point of the Brillouin zone. Medium soft norm-conserving pseudopotentials of the Martins−Trouillier type[79] are used. The core−valence interaction of C, N, and O is treated by s and p potentials with pseudization radii of 1.23, 1.12, and 1.05 au, respectively (taking the same radius for s and p). Energy expectation values are computed in reciprocal space using the Kleinman−Bylander transformation.[80] Hydrogen atoms are represented by a simple s potential damping the Coulomb singularity at the origin. We used the Becke, Lee, Yang, and Parr (BLYP) gradient-corrected functional[81,82] for the exchange and correlation terms.

The simulations were performed at constant volume using a fictitious electron mass of 600 au and a time step of 5 au (0.12 fs). In the equilibration phase, the temperature was controlled by velocity scaling. To avoid any interference with the dynamics, the trajectory over which the data for the spectrum was collected is strictly microcanonical. The average ionic temperature obtained over the length of this run (7.3 ps) was 310 K (the 10 deg in excess of room temperature was not intended but is the result of the microcanonical conditions). Hydrogen atoms were treated as classical particles with their true mass (1836 au).

The 11.5 Å cell size of our model system (see snapshot Figure 1) is a compromise between spatial extent and duration of the MD run. The starting configuration was prepared using classical force field methods. This included the determination of the precise number (49) of water molecules in the cell. A similar trial and error procedure for fixing the density was used in ref 21. Details can be found in ref 71. The length of the trajectory is particularly critical in real time spectroscopic applications because it determines the spectral resolution. The total duration of the data collection run, 7.3 ps, should be sufficient to achieve a resolution on the order of $10-20$ cm^{-1}, which is the minimum accuracy of interest to vibrational spectroscopy in the condensed phase. Computation of the dipoles of the molecules, that is, determination of Wannier functions, was carried out every third step. The cost in CPU time of a step including this operation is approximately 5 times more than a simple dynamics step.

We have also performed a Car−Parrinello molecular dynamics simulation of a uracil molecule in a empty (periodic) cubic box of 12 Å at the temperature of 20 K. The infrared spectrum obtained from this gas-phase calculation will be tested against the frequencies computed from second derivative methods and will serve as a reference for analyzing solvent effects. The

Feature Article

J. Phys. Chem. B, Vol. 107, No. 38, 2003 **10349**

TABLE 1: Geometrical Parameters (Distances in Å, Angles in deg) of Uracil

	bond lengths						bond angles				
bond	PW (0 K) vacuum[a]	AE (0 K) vacuum[b]	MD (20 K) vacuum[c]	expt[d]	MD (310 K) solution[e]	angle	PW (0 K) vacuum[a]	AE (0 K) vacuum[b]	MD (20 K) vacuum[c]	expt[d]	MD (310 K) solution[e]
N1−H1	1.020	1.020	1.020 ± 0.007	0.836 (1.002)	1.042 ± 0.035	C2−N1−C6	123.6	123.8	123.5 ± 0.5	122.7 (123.2)	122.3 ± 2.2
N1−C2	1.406	1.411	1.406 ± 0.007	1.371 (1.399)	1.386 ± 0.028	C6−N1−H1	121.2	121.2	121.1±0.8	112.1 (121.0)	120.4 ± 4.1
C2=O2	1.229	1.234	1.229 ± 0.002	1.215 (1.212)	1.247±0.021	H1−N1−C2	115.1	115.0	115.1 ± 0.7	115.1 (115.1)	115.9 ± 4.1
C2−N3	1.392	1.398	1.393 ± 0.008	1.377 (1.399)	1.386±0.027	N1−C2=O2	122.7	122.7	122.8 ± 0.9	123.7 (123.8)	122.4±2.6
N3−H3	1.023	1.023	1.024 ± 0.001	0.877 (1.002)	1.042±0.032	N1−C2−N3	112.8	112.7	112.7 ± 0.6	114.0 (114.0)	114.3±1.9
N3−C4	1.427	1.431	1.428 ± 0.009	1.371 (1.399)	1.405±0.029	O2=C2−N3	124.5	124.6	124.4 ± 0.7	122.3 (121.9)	122.5 ± 2.4
C4=O4	1.232	1.237	1.233 ± 0.007	1.245 (1.212)	1.258 ± 0.022	C2−N3−H3	115.7	115.6	115.7 ± 0.7	117.8 (129.3)	116.4 ± 3.3
C4−C5	1.459	1.466	1.460 ± 0.008	1.430 (1.462)	1.441 ± 0.026	C2−N3−C4	128.0	128.1	127.7 ± 0.7	126.7 (126.0)	125.9 ± 2.2
C5=C6	1.353	1.365	1.354 ± 0.005	1.340 (1.343)	1.360 ± 0.021	H3−N3−C4	116.2	116.2	116.1 ± 0.8	115.5 (116.1)	115.8 ± 3.4
C5−H5	1.085	1.089	1.085 ± 0.005	0.970 (1.072)	1.085 ± 0.018	N3−C4=O4	120.0	120.1	119.9 ± 0.7	119.2 (120.2)	118.7 ± 2.4
C6−H6	1.088	1.092	1.088 ± 0.004	0.957 (1.072)	1.089 ± 0.027	N3−C4−C5	113.5	113.5	113.4 ± 0.5	115.5 (115.5)	114.5 ± 2.3
C6−N1	1.382	1.387	1.382 ± 0.004	1.359 (1.399)	1.375 ± 0.025	O4=C4−C5	126.5	126.4	126.6 ± 0.8	125.3 (124.3)	126.2 ± 3.1
						C4−C5−H5	118.2	118.1	118.1 ± 0.9	118.1 (118.1)	118.6 ± 4.2
						C4−C5=C6	120.2	120.1	120.1 ± 0.7	118.9 (119.7)	119.1 ± 2.2
						H5−C5=C6	121.6	121.7	121.6 ± 0.7	123.0 (122.5)	121.4 ± 4.3
						C5=C6−H6	122.7	123.0	122.7 ± 0.7	123.2 (122.8)	122.4 ± 3.9
						C5=C6−N1	122.0	121.7	121.9 ± 0.6	122.3 (122.1)	121.8 ± 2.4
						H6−C6−N1	115.3	115.2	115.3 ± 0.8	114.5 (115.1)	115.1 ± 3.4

[a] Geometry optimization using BLYP and the plane wave and pseudopotential representation applied in Car−Parrinello calculations. [b] Full electron geometry optimization at the BLYP/6-31++G* level. [c] Average values and root-mean-square deviations from a Car−Parrinello MD simulation of isolated uracil (20 K). [d] Experimental values from X-ray (ref 90) and electron diffraction data (ref 91, in parentheses). [e] Average values and root-mean-square deviations from MD simulation of aqueous uracil (310 K).

**TABLE 2: Harmonic Vibrational Frequencies (cm^{-1}) for the 0 K Equilibrium Geometries of Table 1**

PW[a]	AE[b]	PW[a]	AE[b]	PW[a]	AE[b]	PW[a]	AE[b]
3479	3523	1335	1382	875	922	464	519
3444	3483	1327	1360	771	785	440	503
3154	3183	1313	1331	742	741	379	383
3120	3143	1190	1203	711	705	306	368
1726	1727	1151	1142	683	698	156	162
1694	1689	1032	1053	589	666	93	141
1599	1611	937	957	556	551		
1420	1451	926	929	472	536		

[a] Plane wave and pseudopotential calculation. [b] Gaussian BLYP/6-31++G* calculation.

temperature of 20 K was chosen, close to the thermodynamic conditions of the infrared spectroscopic experiment of uracil in an argon matrix,[83] to which we will compare our theoretical investigation. The initial configuration of the DFT-MD run was the optimized geometry of uracil. The duration of simulation was 6.5 ps. All simulations were carried out with version 3.4.1 of the CPMD ab initio molecular dynamic package.[84]

3. Uracil in Vacuo

Equilibrium Geometry and Vibrations. The accuracy and efficiency of the electronic structure calculation scheme applied in the Car−Parrinello method depend on a careful balance between the size of the plane wave basis set and the level of pseudization ("softness") of the core pseudopotentials. Numerous applications (see papers quoted in section 1) have shown that the standard combination of medium-soft Trouillier−Martins pseudopotentials and a 70 Ry plane wave cutoff (section 2) is adequate for the description of most equilibrium properties of aqueous systems under ambient conditions. To verify that this electronic representation also meets the requirements for accurate computation of (harmonic) vibrational frequencies, the geometry of isolated uracil was optimized using standard gradient techniques followed by numerical determination of second derivatives. The geometry and frequencies were compared to the results of a Gaussian[85] all-electron optimization at the BLYP/6-31++G* level (calculation realized in this work). Results are reported in Table 1 (geometries) and Table 2 (frequencies).
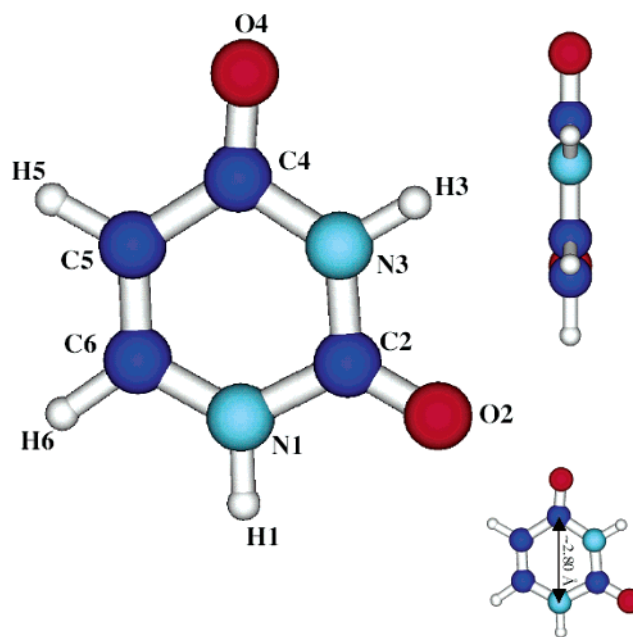
Figure 2. Optimized geometry of uracil obtained in this work with the labeling of atoms used in the text (values for bond lengths and angles are given in Table 1).

Both geometry optimizations give a planar structure for uracil (C_s symmetry, Figure 2). The plane wave−pseudopotential estimations of bond lengths are ∼0.005 Å shorter compared to the all-electron calculation. The exception is the C=C bond length, which was found to be 0.012 Å longer in the plane wave calculation. Bond angles are identical. Moreover, all values are very similar to reference data from the literature.[86−89] In agreement with the general trend for this generalized gradient approximation (GGA) functional, the bond lengths obtained using BLYP/6-31++G* are 1% longer compared to the B3LYP/6-31++G* values.[88] Even so, we observe an overall satisfactory agreement with X-ray and electron diffraction measurements.[90,91] Vibrational wavenumbers are listed in Table 2. The frequencies from the pseudopotential calculation are ∼10−50 cm^{-1} lower relative to the all-electron calculation. Consistent with the overestimation of bond length, the vibrational wavenumbers
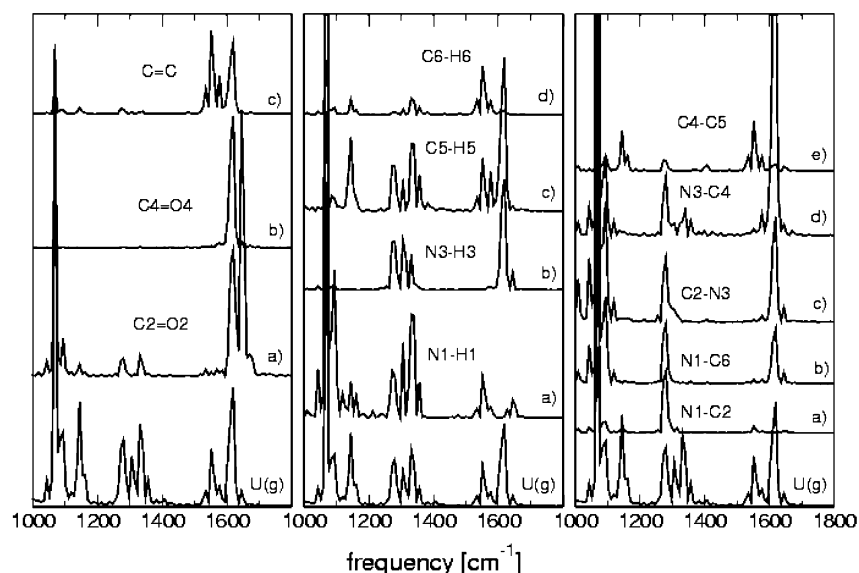
Figure 3. Vibrational density of states (VDOS) in the 2000−1000 cm⁻¹ spectral domain of isolated uracil. Panels show the decomposition according to internal movements (see section 3 for description of the method used). The bottom spectrum is the same in all three panels and shows the full VDOS as obtained from a 6.5 ps Car−Parrinello molecular dynamics trajectory at a temperature of 20 K. The left panel shows the stretching vibrational motions: (a) C2═O2; (b) C4═O4; (c) C═C. The middle panel shows the bending vibrational motions: (a) N1−H1; (b) N3−H3; (c) C5−H5; (d) C6−H6. The right panel shows the ring stretching vibrational motions: (a) N1−C2; (b) N1−C6; (c) C2−N3; (d) N3−C4; (e) C4−C5.

TABLE 3. Comparison of Calculated Vibrational Density of States of Uracil in a Vacuum and Experimental Infrared Spectrum of Uracil in an Argon Matrix

MD[a]	assignments[c]	expt[b]	MD[a]	assignments[c]	expt[b]
3340	ν(N1−H1)	3485−3474	1356	δ(N1−H1); δ(C−H)	1472−1459
3298	ν(N3−H3)	3435	1331	δ(N−H); δ(C−H); ν(C═O)	
3026	ν(C5−H5)		1305	δ(N−H); δ(C−H)	
2984	ν(C6−H6)		1280	δ(N−H); δ(C−H); ν(ring)	1399−1387
1645	ν(C═O)	1792−1681	1161, 1144	δ(C−H); ν(C═C)	
1619	ν(C═O)	1792−1681	1094, 1068, 1043	ν(ring); δ(ring)	1217−1176
1576, 1560, 1534	ν(C═C); δ(C−H)	1644			

[a] Band positions (cm⁻¹) in the vibrational density of states (Figure 3) computed from the 20 K trajectory of isolated uracil. *[b]* Experimental infrared spectrum of uracil in argon matrix (ref 83). *[c]* Notations: ν (bond-stretch), δ (bend, angle deformation). Assignments according to the method described in section 3.

obtained from the all-electron BLYP/6-31++G* calculation are down-shifted by up to ~5−10% relative to our previously published B3LYP/6-31++G* frequencies.[88,89]

**Low-Temperature Geometry and Vibrational Dynamics.** A crucial question for a dynamical approach to spectroscopy is how the properties estimated by averaging over a finite temperature dynamical trajectory compare to the values obtained from perturbation expansions at potential energy surface minima. The motion of an isolated uracil molecule (see section 2 for technical details) consists at 20 K of small distortions of an on-average planar ring with amide, carbonyl, and C−H extraring groups vibrating in and out of the plane. The average geometry parameters can be found in Table 1, in which we can see that they are virtually identical to the optimized structure obtained using the same electronic representation. As expected at these low temperatures, root-mean-square deviations are very small.

The decomposition of the corresponding VDOS in internal motions is shown in Figure 3. This decomposition is the basis for the assignments given in Table 3. Depending on the spectral region of analysis, we find that the peaks displayed by the VDOS are red-shifted by 40−140 cm⁻¹ with respect to the corresponding harmonic frequencies computed at 0 K. High-frequency modes, that is, N−H and C−H stretches, are most affected. The temperature (20 K) is too low to give rise to anharmonic effects of a magnitude that could explain such down shifts. Failure to recover the harmonic limit may be due to an

artifact in the Car−Parrinello method, namely, coupling to the classical inertia of the electrons. If true, this should be manifested in a dependence on the choice of the parameter associated with the fictitious electron mass. Accordingly, we have checked this hypothesis by varying the value of this parameter, decreasing it to 300 au, and running a short dynamics trajectory monitoring the highest N−H and C−H stretching frequencies for which the deviations were largest. We found that these values are increased and become 3447, 3391, 3108, and 3052 cm⁻¹, respectively for ν(N1−H1), ν(N3−H3), ν(C5−H5), and ν(C6−H6), representing a shift of 40−50 cm⁻¹ from the corresponding harmonic frequencies (for the atomic labeling, see Figure 2). The appreciable underestimation by 120−140 cm⁻¹ of the frequencies of the maxima in the VDOS in comparison to the values measured in experiments[83] (data reproduced in Table 3) is therefore a combination of electronic inertial effects and softening due to the density functional. Fortunately, whereas band positions obtained from the calculated vibrational density of states are down-shifted with respect to experiment, the spectral patterns displayed in the calculation are consistent with observation.

Assignment of Vibrational Modes. Assignment to internal movements, as described in section 2, gives the following main results (see Figure 3 and Table 3). The spectral patterns in the 1700−1500 cm⁻¹ range are due to C═O and C═C bond vibration. The carbonyl groups are mainly responsible for the

Feature Article

J. Phys. Chem. B, Vol. 107, No. 38, 2003 **10351**
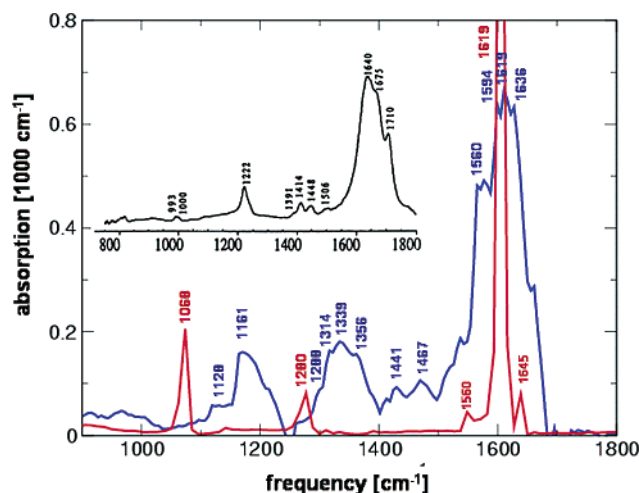
Figure 4. Infrared absorption spectrum of uracil in vacuo (red) and in aqueous solution (blue) as computed from Car−Parrinello trajectories at a temperature of 20 and 310 K, respectively. Displayed is the 2000−1000 cm^{-1} spectral domain of interest in biochemistry without any quantum corrections applied (see text and appendix for results on detailed balance corrections). The intensity of the 1619 cm^{-1} vacuum band is 6800 cm^{-1} and has been truncated in the graph. Inset shows the experimental infrared spectrum of aqueous uracil taken from ref 69.

bands at 1645 and 1619 cm^{-1}. The mode at 1645 cm^{-1} is dominated by C2=O2 stretching, while the mode at 1619 cm^{-1} results from the superposition of C2=O2 and C4=O4 stretching. The three bands between 1576 and 1534 cm^{-1} are related to C=C stretching.

Bending involving N−H or C−H bonds induces vibration of adjacent C=O and C=C bonds, which explains the appearance of signatures of bending movements in the 1700−1500 cm^{-1} stretching regime. The origin of the four bands in the 1400−1200 cm^{-1} range is predominantly δ(N−H) and δ(C−H) bending. The band at 1356 cm^{-1} is due to δ(N1−H1). Clearly, bending of the other amide group (N3−H3) is not participating in this mode, while some coupling to δ(C−H) can be observed. The three other bands, located at 1331, 1305, and 1280 cm^{-1}, are related to a superposition of δ(N−H) and δ(C−H) bending. We note that ν(ring) stretching is an important component of the vibrational band at 1280 cm^{-1}. In addition to in-plane bending displacements, this mode induces strong deformations of the uracil ring, capable of changing the molecular dipole. This suggests a much higher IR activity for this mode in comparison to the other modes in the 1400−1200 cm^{-1} range, which are not accompanied by such ring deformations. The two bands located at 1161 and 1144 cm^{-1} are assigned to pure C−H bending motions. Bands between 1100 and 1000 cm^{-1} are due to an overall superposition of ν(ring) and δ(ring) ring stretching and bending.

Infrared Spectrum. The same trajectory of 6.5 ps that was the source of the data for the VDOS analysis (Figure 3) was also used to compute the infrared absorption spectrum according to the method described in section 2. The bare spectrum, prior to application of any quantum correction, is reported in Figure 4 (the discussion of quantum effects is deferred to the Appendix). Figures 5 and 6 contain the same data but now with a graphical representation of the assignments as obtained from the VDOS superimposed. In the spectral domain between 1700 and 1500 cm^{-1}, three active infrared bands can be distinguished. The most intense peak located at 1619 cm^{-1} is very narrow. The two other bands, at 1645 and 1560 cm^{-1}, have intensities about 20−40 times lower. The 1619 cm^{-1} band is due to
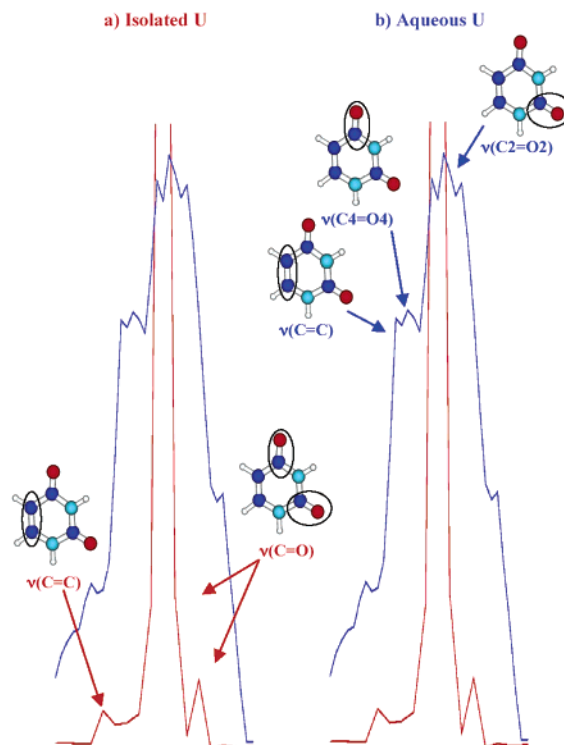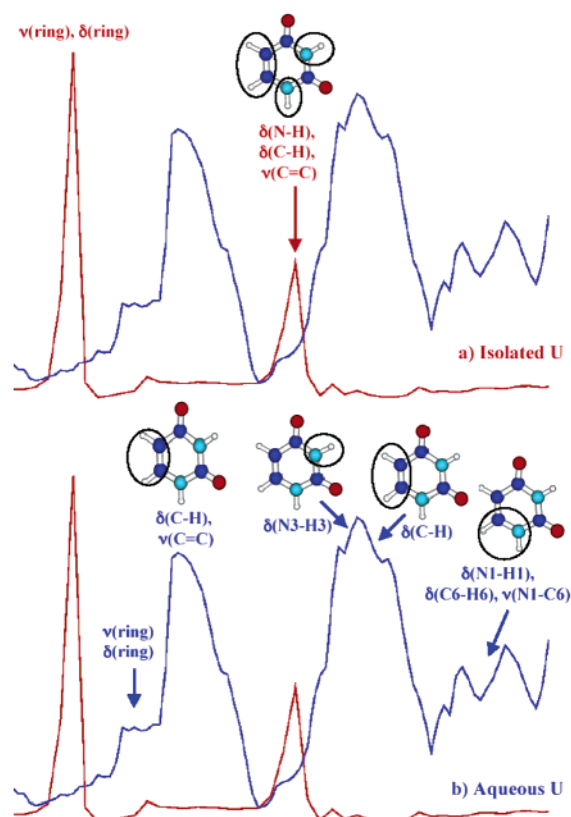

Figure 5. Schematic representation of the assignments of the infrared active vibrational bands of (a) isolated and (b) aqueous uracil in the C=O stretch domain (1700−1500 cm^{-1}; for full spectrum, see Figure 4).

carbonyl stretching ν(C=O). This vibration is the most effective for modulation of the molecular dipole, hence the high intensity. The band at 1645 cm^{-1} is also related to carbonyl stretching motion, but consisting of almost pure C2=O2 vibration (see Figure 3), its infrared activity is reduced relative to the 1619 cm^{-1} band in which both carbonyl vibrations participate. The infrared band at 1560 cm^{-1} is due to the stretching motion ν(C=C) of the C=C double bond with some coupling to δ(C−H) bending.

Two absorption peaks were found in the spectral domain between 1500 and 1000 cm^{-1} of medium intensity (1280 cm^{-1}) and high intensity (1068 cm^{-1}). These intensities, however, are far lower than the activity of the carbonyl stretching mode. As we saw in the VDOS analysis, the band located at 1280 cm^{-1} is a superposition of δ(N−H) and δ(C−H) bending and ν(ring) ring stretching. Recall that in the same spectral domain three other vibrational bands were observed in the VDOS (Figure 3), which were also related to amide and C−H bending motions. However, as anticipated, the only IR active band is the one involving ring stretching. Similarly, the infrared band at 1068 cm^{-1}, associated with diverse deformations of uracil ring (ν(ring) ring stretching and δ(ring) ring bending), also carries high intensity.

One to one comparison of the calculated infrared bands of isolated uracil to the experimental spectrum recorded in argon matrixes[83] in the relevant 1800−1500 cm^{-1} spectral domain is unfortunately not possible. The complication is Fermi interaction, which, as explained in ref 83, leads to splitting of the two carbonyl stretching modes (fundamentals) into several peaks (combination bands). Fermi resonances are not well described in our classical approach, and therefore, the positions and number of bands observed in our calculation have no clear correlation to the different bands observed in the experiment. Nonetheless, experiment suggests a gap of ~100 cm^{-1} between

10352 *J. Phys. Chem. B, Vol. 107, No. 38, 2003*

Gaigeot and Sprik

Figure 6. Schematic representation of the assignments of the infrared active vibrational bands of (a) isolated and (b) aqueous uracil in the C−N−H bending domain (1500−1000 cm^{-1}; for full spectrum, see Figure 4).

the zone related to C=O and C=C stretches, which is consistent with our calculations (Figure 4, Table 3). The C=O infrared band has the highest intensity in the calculated spectrum, again in qualitative agreement with experiment. In the spectral range between 1500 and 1000 cm^{-1}, the experimental infrared spectrum is composed of three active bands, located in the intervals 1472−1459 cm^{-1} (small intensity), 1399−1387 cm^{-1} (medium intensity), and 1217−1176 cm^{-1} (high intensity). Only two infrared bands are predicted in our calculation. In Table 3, we have assumed that the absorption peak at 1280 cm^{-1} (medium intensity) corresponds to the band measured at 1399−1387 cm^{-1}, while the 1068 cm^{-1} (high intensity) band should be compared to 1217−1176 cm^{-1} band in the experiment. Despite the down shift with respect to the experiment, the differences in frequencies between these two bands are 182−223 and 212 cm^{-1}, respectively, in the experiment and in our calculation, which is a very good agreement. Responsible for the overall red shift in the calculation can be the BLYP functional, as well as the value of the fictitious electron mass used in the Car−Parrinello dynamics. This is probably not the reason the vibrational band appearing at 1356 cm^{-1} in the VDOS (Figure 3, Table 3) is missing from the computed IR spectrum. This band, related to N1−H1 bending motion, is observed with a small intensity in the experiment. The reason the calculation does not give any infrared activity to this mode could be an underestimation of the dynamical charges associated with this amide group.

4. Uracil in Solution

Hydrogen Bonding. The broad picture of the hydration of uracil emerging from the ab initio MD simulation[71] conforms to what is known or can be anticipated for this class of aqueous molecules. Amide and carbonyl groups participate fully in solute−water intermolecular hydrogen bonding. Bond lengths show the regular variation from 1.70, 1.80 Å to 2.30, 2.50 Å, measured from the first peak in the radial distribution to the next minimum, leading to one bond per amide H and two bonds per carbonyl O atom. The C5−H5 group is not involved in intermolecular hydrogen bonds, and the closest water molecules to this site belong to the second hydration shell of the solute. The other CH group, however, C6−H6, showed a hint of hydrogen bonding with very long (2.40−2.50 Å), distorted (110°−120°), and short-lived bonds. The average number of solvent molecules hydrogen bonded to the main hydrophilic sites of uracil is therefore six. In addition, the full first hydration shell was found to contain a maximum of three water molecules positioned in van der Waals contact above and below the uracil ring. Also this predominantly hydrophobic coordination shows some directional preference suggesting a weak attraction of the water H atoms to the electronic π cloud (see ref 71 for a detailed analysis).

Intramolecular Structure. Table 1 also lists some structural details of uracil as computed by averaging over the DFT-MD simulation of the aqueous system. These structural parameters are compared to the values from the Car−Parrinello molecular dynamics of the isolated molecule (see also discussion of section 3). Due to the higher temperature in the dynamics of aqueous model system (310 K), variations around the mean values are more prominent than those for the colder molecule in a vacuum (20 K). Solvation induces a number of changes in internal geometry. Sites that form hydrogen bonds with the surrounding water molecules are most affected. Amide bonds are stretched by +0.022 and +0.018 Å, respectively, for N1−H1 and N3−H3 and carbonyl bonds by +0.018 and +0.025 Å for C2=O2 and C4=O4, respectively. The 0.011 Å larger bond length for C4=O4 compared to C2=O2 is probably significant in view of the slightly higher number of water molecules coordinated to C4=O4, as well as the tighter H-bonds to this site as revealed by the analysis of ref 71. The C5−H5 distance is maintained at the isolated molecule value as expected because this group is not involved in hydrogen bonding. Moreover, despite evidence of some C6−H6···O$_w$ attraction, the weakness of this bond does not lead to any modifications of C6−H6 bond length. On the other hand, a substantial overall contraction of the uracil ring is observed upon solvation. N−C and C−C bond lengths are reduced by 0.020 Å. An exception is the C=C bond length, which is increased by a small amount (0.006 Å).

5. Static Polarization

Dipole Moment of Aqueous Uracil. In preparation of the discussion of the results of the calculations of the infrared absorption, we now look at the static polarization of uracil in aqueous solution as calculated using the maximally localized Wannier function approach outlined in section 2. The solvent has a significant effect. We find that the mean value of the dipole is 7.46 D. The root-mean-square deviation from the mean value is 0.74 D (~10%). This should be compared to the 4.50 D obtained by the same method in the gas phase. Our 4.50 D is in very good agreement with calculations from ref 87 and slightly higher than the experimental value.[92]

Responsible for the ~65% enhancement of the dipole moment is most likely the rather asymmetrical (lopsided) configuration of the nine molecules in the first solvation shell,[71] cooperatively pulling the electrons in one direction. During rare fluctuations, the instantaneous (adiabatic) dipole moment can take values as

Feature Article

J. Phys. Chem. B, Vol. 107, No. 38, 2003 **10353**

low as 5.45 D, which is only a 1.00 D increase over the value of the isolated molecule, and values as high as 10.10 D, which is more than twice the dipole value of the isolated molecule. The direction of the dipole moment of uracil, in contrast, does not change much from the gas to the aqueous phase. Isolated uracil has a dipole moment that is contained in the plane of the molecule; it is roughly directed along the N3 to C6 axis and points away from the side with the two carbonyl groups. When solvated by water, the uracil dipole remains in the plane of the molecule (the vector acquires a very small component out of the plane of uracil), and its direction oscillates on average by $10° ± 5°$ around the isolated molecule reference.

Dipole Moment of Solvent Molecules. The tetrahedrally coordinated hydrogen bonding in liquid water strongly enhances the molecular dipole moments.[5] It is therefore an interesting question whether the presence of the solute disrupts this cooperative effect. We found that, apart from a small increase of 0.1−0.2 D, the dipole moments of the water molecules making up the first hydration shell have essentially the same distribution of values compared to the bulk solvent. The most probable value measured over all 49 molecules is again 2.90−3.00 D, identical to the results of Silvestrelli and Parrinello[5] for pure liquid water. Instantaneous dipoles vary between 2.00 and 4.20 D. Perturbations by the solute are therefore small. This is consistent with another result of the structural analysis of ref 71, namely, the lack of modification of internal geometries of the water molecules. So, evidently, while the uracil is highly polarized in solution, this interaction is not reciprocal.

6. Infrared Spectrum in Solution

Solvent Spectrum. Before analyzing the details of the solute spectrum, we first consider the infrared absorption spectrum of the water solvent as calculated from the MD trajectory of the solution (Figure 7). Of the 49 solvent molecules in the model only vibrations of the nine water molecules that belong to the first hydration shell are directly affected by intermolecular H-bonds formed with the solute. This number of water molecules is too small in comparison to the total number of water molecules in the box to cause drastic shifts or variations in intensity of the vibrational bands of water with respect to pure liquid. Moreover, as noted in section 5, the dipole moments of both ligand and bulk solvent molecules are almost insensitive to interactions with the nucleic acid base. We can therefore expect equally little change in the infrared absorption of water solvent compared to the pure liquid.

As can be seen in Figure 7, the calculated infrared spectrum of the water solvent is characterized by four major bands located at 220 cm^{-1} (intermolecular H-bond stretching[49]), 639 cm^{-1} (librational peak[49]), 1611 cm^{-1} (intramolecular bending), and 3069 cm^{-1} (O−H stretching). For a comparison, the experimental infrared spectrum of pure liquid water[93,94] is also reproduced in Figure 7. Positions of absorption bands obtained in our calculation are in very good agreement, apart from the OH stretching band which, similar to the N−H and C−H bands of uracil, is down-shifted by ∼250 cm^{-1}. The reasons for this discrepancy are also the same and have been discussed at length in section 3. Again for technical comments regarding quantum corrections, we refer to the Appendix.

Solute Spectrum: General. Solvent shifts and other solvation effects are best discussed when compared to a reference vacuum spectrum determined using a computational methodology as similar as possible. Hence, our reference will be the absorption spectrum obtained from the 20 K ab initio MD run in an empty (periodic) MD box of 12.0 Å. Both the vacuum
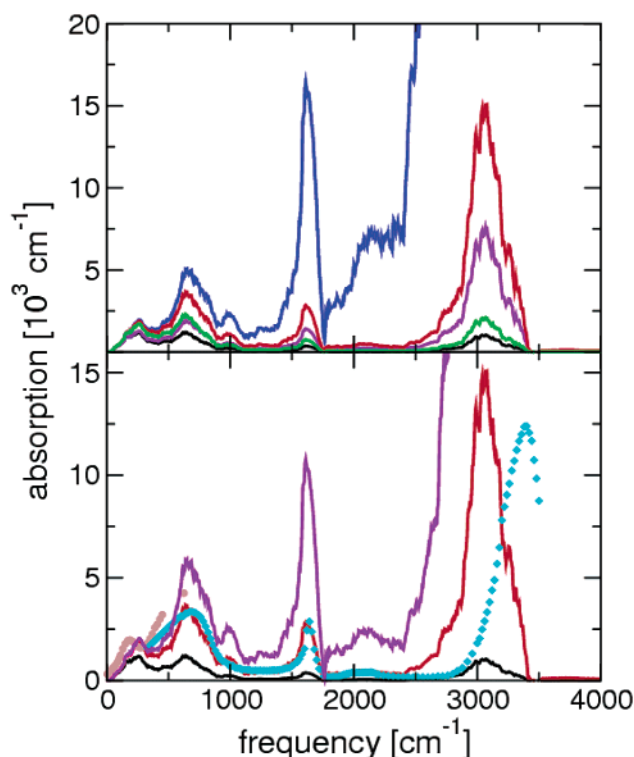

Figure 7. Infrared absorption spectrum of water solvent computed from the dynamics trajectory of aqueous uracil. The results of the application of a variety of quantum and detailed balance corrections are also displayed (see Appendix). The top panel shows the spectra calculated without quantum correction (black), with $2/(1 + \exp(-\beta\hbar\omega))$ correction (green), with $\beta\hbar\omega/(1 - \exp(-\beta\hbar\omega))$ correction (red), with $\exp(\beta\hbar\omega/2)$ correction (blue), and with the mode amplitude correction $\beta\hbar\omega/(2 \tanh(\beta\hbar\omega/2))$ (violet). The bottom panel show the calculated infrared spectrum of water solvent compared to the experimental infrared spectrum of liquid water. Experimental points represented by circles are taken from ref 94, and those represented by diamonds come from ref 93. The calculated spectrum without quantum corrections is reported in black, the corrected one with the $\beta\hbar\omega/(1 - \exp(-\beta\hbar\omega))$ factor in red, and the corrected one with the combination of $\beta\hbar\omega/(1 - \exp(-\beta\hbar\omega))$ and mode amplitude term in violet.

and aqueous spectrum were given in Figure 4. Again we have tried to match the vibrational signatures of local modes with the total VDOS in solution (Figure 8). The resulting interpretation of the spectrum is summarized in Figures 5 and 6 and compared to the vacuum assignments.

Two strikingly contrasting features are apparent. On one hand, solvation has little effect on the spectrum between 1700 and 1500 cm^{-1}. Broadening seems the main response in this interval. On the other hand, the number and positions of infrared active bands is completely different in the 1500−1000 cm^{-1} spectral range. Indeed, the IR absorption in a vacuum is characterized by only two infrared bands in this domain, while five (of various intensities) can be distinguished in solution. Because of the condensed-phase environment, most of the calculated bands are broader than those in the gas phase. The order of magnitude of the intensities, however, remains similar.

CO and CC Stretching. A broad intense infrared band can be observed extending from 1700 to 1500 cm^{-1}. The intramolecular motion with the best correlation to this part of the spectrum is C=O and C=C stretching (Figure 8), as was the case for isolated molecule in the same domain (Figure 3). Stretching of the C2=O2 carbonyl group mainly shows up in the higher frequency range of this broad band, that is, for wavenumbers above 1594 cm^{-1}. The C4=O4 and C=C
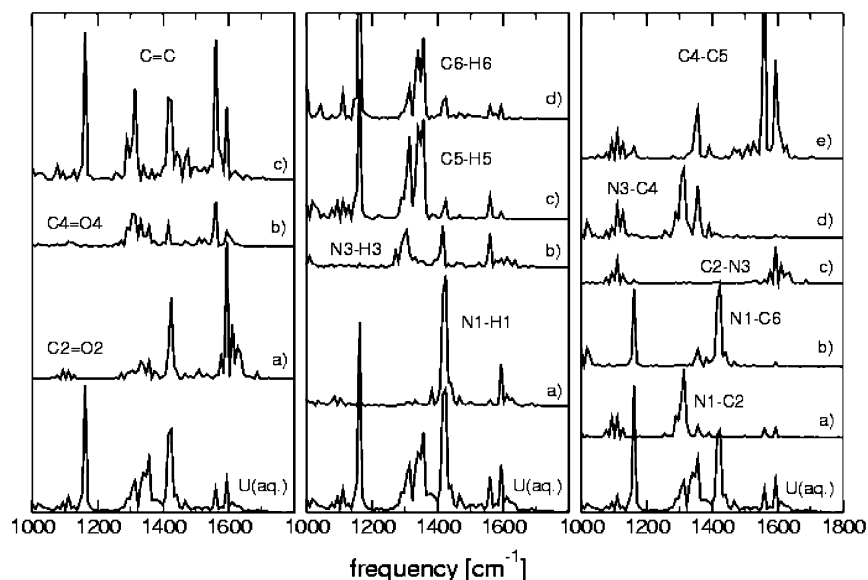
Figure 8. Decomposition of the VDOS of aqueous uracil according to internal movements. The full VDOS is repeated at the bottom of each panel. The left panel shows the stretching vibrations: (a) C2=O2; (b) C4=O4; (c) C=C. The middle panel shows the bending motions: (a) N1−H1; (b) N3−H3; (c) C5−H5; (d) C6−H6. The right panel shows the ring stretching motions: (a) N1−C2; (b) N1−C6; (c) C2−N3; (d) N3−C4; (e) C4−C5 (compare Figure 3).

stretching modes merge in the low-frequency wing of the band at 1560 cm^{-1}. Carbonyl C4=O4 group bond vibration also contributes to the 1594 cm^{-1} band. As already outlined for the case of isolated uracil, carbonyl and C=C stretching is accompanied by N−H and C−H bending motions, as revealed by signatures of these movements in the decomposition of the VDOS of the aqueous species.

NH and CH Bending. The 1500−1000 cm^{-1} spectral domain can be divided into five infrared active bands (see Figure 4 for data and Figure 6 for a graphical interpretation). Three have medium intensities (1467, 1441, and 1128 cm^{-1}), and two have higher intensities (1339 and 1161 cm^{-1}). All intensities are lower than the intensity of the main carbonyl-dominated band discussed above. The infrared bands between 1400 and 1100 cm^{-1} are relatively broad, whereas each of the bands in the 1500−1400 cm^{-1} range is very narrow. We note also that the band marked at 1441 cm^{-1} shows a tail (or shoulder) at 1420 cm^{-1}. It is difficult to determine whether this feature is meaningful. We will discuss this later, when comparing our calculated spectrum to experiment. Going further down in frequency to the broad band between 1400 and 1300 cm^{-1}, a number of finer structures can be distinguished at 1356, 1339, 1314, and 1288 cm^{-1}. We will see below that they can be assigned to specific intramolecular atomic displacements.

When we try to analyze the spectrum in the 1500−1000 cm^{-1} domain, significant differences appear with respect to isolated uracil. The main points concern a splitting among the amide bending modes in combination with the appearance of an infrared active mode associated with C−H bending. The VDOS analysis (Figure 8) suggests that N3−H3 bending gives rise to an infrared active band down-shifted from the N1−H1 bending and is mixed in with C−H bending and some N1−C6 stretching. This leads us to interpret the vibrations at 1467 and 1441 cm^{-1} as exclusively δ(N1−H1) bending motion with no contributions from N3−H3 bending. The origin of the narrow tail at 1420 cm^{-1} is also mainly amide bending but now with N1−H1 and N3−H3 both represented. The next infrared active band, located between 1400 and 1300 cm^{-1}, can be assigned to δ(N3−H3) and δ(C−H) motions. The higher wavenumbers, that is, 1356 and 1339 cm^{-1}, can be traced back to pure C−H bending

motion, the lower part of the band, that is, wavenumbers 1314 and 1288 cm^{-1}, is associated with a superposition of these two bending motions. Again signatures of N−C and even C=C bond stretches can be detected. N1−H1 bending, however, is absent from this part of the spectrum. Finally at the low-frequency end, the intense band at 1161 cm^{-1} can be assigned to δ(C−H) bending and C=C stretching motions with a small ν(C6−N1) component.

Analysis of Solvent Effects. The effect of hydration on the carbonyl stretching modes is found to be relatively straightforward and common to many aqueous organic compounds. The carbonyl frequencies are red-shifted by ∼10−40 cm^{-1}. Such red shifts are consistent with the formation of hydrogen bonds between water and the carbonyl groups. The C=C stretch is not affected by hydration because the infrared band of this mode is located at 1560 cm^{-1} in the isolated and aqueous molecule. This is in agreement with the more hydrophobic nature of this side of the molecule.

In contrast, the response of the bending modes is surprisingly complex. The analysis of the bending above leads to the conclusion that the amide bending band at 1280 cm^{-1} in the infrared spectrum of isolated uracil splits into two distinct multiplets: two bands related to N1−H1 bending (1467 and 1441 cm^{-1}) and two bands to N3−H3 bending (1314 and 1288 cm^{-1}). Intensities are more or less equivalent to isolated uracil. Solvation, therefore, induces appreciable blue shift in the amide bending motion of uracil (∼20−190 cm^{-1}). The maximum up shift is obtained for N1−H1 bending. Similar blue shifts were also observed in the case of aqueous NMA[48] and seem to be characteristic of amide groups.

Another important observation is that hydration strongly enhances the infrared activity of δ(C−H) bending. Again referring back to the VDOS analysis, two spectral domains were found to be related to C−H bending motions: the higher part of the broad band between 1300 and 1400 cm^{-1} and the 1161 cm^{-1} band. We recall that, under vacuum conditions, C−H bending was coupled to the N−H bending and, hence, also included in the 1280 cm^{-1} infrared active band. At first glance, we would therefore expect the C−H bending in solution either to be blue-shifted (resulting in the 1339 and 1356 cm^{-1} bands),

Feature Article

J. Phys. Chem. B, Vol. 107, No. 38, 2003 **10355**

or red-shifted (resulting in the 1161 cm^{-1} band). However, when we interpreted the VDOS of isolated molecule, we also saw that the two vibrational bands at 1161 and 1144 cm^{-1} were due to C−H bending motion. This band was not infrared active. Thus, one could argue with some reason, that solvation changes the status of this mode into infrared active without any particular shifts. This could be rationalized by the very weak hydrogen bonding exhibited by the C6−H6 group, which could be just large enough to break the symmetry making the line visible. With regard to solvation effects on δ(ring) and ν(ring) ring vibration, the "shoulder" located at 1128 cm^{-1} in aqueous spectrum can be related to this type of motion, which is thus up-shifted in comparison to its counterpart in the vacuum spectrum (1068 cm^{-1}). Such up shift can be explained by the overall contraction of uracil ring upon solvation.

Comparison to Experiment. Figure 4 compares the calculated infrared spectrum of aqueous uracil to the experimental measurements (inset) of ref 70. The overall patterns are very similar, though subtle differences also catch the attention. The same characteristic regimes in intensity and width show up in experiment and calculation. Thus, going from 2000 to 1000 cm^{-1}, one can see in the experimental spectrum a broad and very intense band, three peaks of smaller amplitude, and one single peak of medium intensity (the peak marked at 1391 cm^{-1} appears as a faint shoulder on the band at 1414 cm^{-1}. Therefore, these two peaks are better considered as part of a broader band, extending between ∼1420 cm^{-1} and ∼1390 cm^{-1}).

Theory and experiment are in remarkable good agreement for the main band generated by C=O and C=C stretching, if we allow for a down shift by ∼50 cm^{-1} in the computed spectrum. The calculation then displays two narrow peaks, located at 1467 and 1441 cm^{-1}, and one broad band around 1330 cm^{-1}. The two first, assigned to N1−H1 bending, can be identified with the two narrow bands marked in experiment at 1506 and 1448 cm^{-1}. They are thus down-shifted by 7−39 cm^{-1} from the experiment. Also these discrepancies are acceptable.

However, the band calculated around 1330 cm^{-1} is too broad in comparison to experiment (its width is about twice the experimental one) and too intense. This band has been mostly assigned to N3−H3 bending motion. This suggests (i) that the dynamical charges of the amide N3−H3 group are overestimated in the simulation or (ii) that the hydrogen bonds formed by this group to the water molecules are too strong. Similarly the infrared tail associated in the calculation with the 1441 cm^{-1} band, which has no counterpart in the experimental spectrum, also contains a N3−H3 bending component, which can, therefore, be taken as another indication that the IR activity of the N3−H3 amide atoms is overestimated.

Considering the general trend of softening of modes in the calculation, the 1161 cm^{-1} band, which we assigned to C−H bending and C=C stretching, can be identified with the band appearing in experiment 50 cm^{-1} higher in frequency. The relative amplitudes are similar. Finally, the tail band at 1128 cm^{-1}, consisting according to the analysis of ring stretches and bends, may be contained in the tail of the 1222 cm^{-1} peak in the experiment. Thus, the spectral patterns observed in the infrared spectrum of aqueous uracil are well reproduced by our calculation with wavenumbers underestimated by 50 cm^{-1}. The agreement for the intensities can be further improved if quantum corrections of the form $\beta\hbar\omega/(1 - \exp(\beta\hbar\omega))$ are taken into account (see Appendix).

7. Summary and Conclusions

Calculation of infrared spectra is a severe test for modeling because these spectra are highly sensitive to the potential energy surface. Generally speaking, positions of infrared bands are a test of atomic interactions, while intensities probe atomic charges and their dependence on molecular geometry. Electron polarization is thus an important issue for infrared spectra calculations. This makes it a topic in computational chemistry for which the Car−Parrinello simulation can offer major advantages, because the electronic polarization is directly available in this technique, contrary to classical simulations in which one must rely on supplementary assumptions. The objective that we have set for the present article was to give an impression of state of the art "first-principle" computational spectroscopy in the form of an application to a small but representative aqueous organic molecule of biochemical importance.

The key computational tool for the calculation of the polarization was the maximally localized Wannier function implementation of the modern theory of polarization in periodic systems. This formalism allows a partitioning of the total cell dipole moment in a solute and solvent contribution. While not entirely rigorous or unique, the separation achieved by maximally localized Wannier functions is adequate for condensed molecular systems. The enhancement of the molecular dipole moment defined this way turned out to be substantial, from the 4.50 D of the isolated molecule to 7.50 D in solution, that is, a 65% increase. On the other hand, the same approach finds no evidence of a reverse effect of the solute on the solvent. Dipole moments, even for molecules in the first hydration shell, were essentially identical to the pure liquid water values.

The spectral patterns displayed by our calculations were found in remarkably good agreement with experiment. The general features of the spectrum in 2000−1000 cm^{-1} domain are typical for this type of organic compound. With decreasing frequency, we encounter a broad very intense band due to C=O and C=C stretching, followed by an intermediate region mainly composed of amide N−H bending, and finally a lower part that is a mixture of C−H bending, C=C stretching, and ring stretching. The contribution that ab initio calculation has made is, in our view, in the interpretation of the intermediate bending region, where the aqueous spectrum shows the largest deviation from the vacuum spectrum. Exploiting the microscopic information available from the simulation, we showed that solvation leads to a splitting of the amide bending modes and appearance of an infrared active mode associated with C−H bending and C=C stretching. The narrow amide bending band in a vacuum is spread out over a 500 cm^{-1} wide interval in solution with a substantially blue-shifted high-frequency end. The highest increase in frequency was found for N1−H1 bending. The usual red shift and broadening of C=O stretching bands, which we also observed, are in comparison a much smaller effect.

Outlook. With this detailed account of an application to a single model system, we hope to have presented convincing evidence of the potential of condensed phase ab initio MD based computational spectroscopy. A number of discrepancies and ambiguities remain however. One of the shortcomings in the computed spectrum is a down shift of up to ∼50 cm^{-1}. The origin of this bias, which becomes worse at higher frequencies, is related to the BLYP density functional and further technical approximations inherent to the ab initio MD methodology. An open question of more fundamental nature concerns the status and validity of our method for separating solute and solvent signal. The computational "filter" applied here is a theorist's device. Cross-correlations, which in principle are included in the subtraction methods used in experiment, were ignored in the current treatment. This could also be the explanation for the conflicting results obtained by application of quantum

corrections (see Appendix). Whereas for pure water an appropriate mix of zero-point motion and detailed balance corrections led to better agreement with experiment, there was little improvement for the intensity distribution of the solute spectrum, possibly because of inconsistencies in the way in which we have partitioned the spectrum. The agreement between the computed intensity profile and experiment is, therefore, perhaps somewhat fortuitous. Resolving these problems could require the development of alternative ways of spectral separation. It will also be clear that the methods that we have used for assignment are rather primitive and a more quantitative approach to the analysis of dynamical spectra is needed.

At this early stage of development of the technique, further applications to representative model systems are needed to establish the possibilities and limitations of the MD approach to computational spectroscopy. Computations involving peptides, of course, come to mind first. Also extensions of the present investigation of aqueous uracil nucleic base can be considered, for example, a calculation of the infrared absorption spectrum of its deuterated forms. The response to deuteration is used by experimentalists as a probe of intermolecular hydrogen bonding to specific hydrophilic sites of a solute and could be used in a similar way in computation to eliminate ambiguities in the assignment.

Calculation of the Raman spectrum is another option. The equivalent molecular dynamics computation of Raman activity amounts to the evaluation of time correlation functions of the polarizability tensor. Polarizability, unlike polarization, is a genuine derivative property and is usually computed using perturbative methods. Recently the necessary extensions of density functional methodology in the CPMD code has been put in place by Putrino, Sebastiani, and Parrinello,[52] and the first applications have appeared, for example, to high-pressure ice.[51] These calculations and related studies of NMR chemical shifts showed that the overhead in computational cost of the computation of derivative properties is appreciable but manageable on modern computational hardware platforms. We therefore anticipate rapid progress in this new addition to the computational chemistry of condensed molecular systems.

Appendix: Quantum Corrections

Following Borysow et al.,[95] common choices of the desymmetrization factor (detailed balance correction in the harmonic oscillator approximation) are $2/(1 + \exp(-\beta\hbar\omega))$, $\{\beta\hbar\omega/(1 - \exp(-\beta\hbar\omega))$, and $\exp(\beta\hbar\omega/2)$. A more complex correction, due to Egelstaff (see ref 95), can also be applied as a shift of time origin followed by the $\exp(\beta\hbar\omega/2)$ correction. This was the factor used by Silvestrelli et al. in their water calculation.[49] For thermal conditions and the relatively high intramolecular vibrational frequencies (>600 cm^{-1}) of interest here, exponential

growth is the dominant part of the Egelstaff correction, which, in this regime, reduces to the simple third expression for detailed balance given above. Moreover, the validity of the Egelstaff procedure at frequencies above 600 cm^{-1} has often been questioned.[95] This correction was therefore not considered any further. The question of quantum corrections is further complicated by the amplitude corrections necessary to take into account hydrogen atom zero-point motion effects. This factor has the form $\beta\hbar\omega/(2 \tanh(\beta\hbar\omega/2))$ and can be applied, alone or combined with desymmetrization (for example, note that combining the $2/(1 + \exp(-\beta\hbar\omega))$ detailed balance factor with the mode amplitude correction is identical to the $\beta\hbar\omega/(1 - \exp(-\beta\hbar\omega))$ scheme).

There clearly is a certain degree of arbitrariness in the choice of the quantum correction. We thus will examine the influence of various combinations of corrections on the calculated infrared spectra, and we will try to show that certain choices give more reasonable results for the systems investigated here.

Solvent Spectrum. Figure 7 compares the results of the application of quantum correction to the calculated classical line shape. Frequency values are invariant under the detailed balance correction, but we can clearly see different behavior of the band intensities. The general effect is a desymmetrization of the amplitudes of the two lower bands (220 and 639 cm^{-1}). In particular, the amplitude of the librational band is increased in comparison to the 220 cm^{-1} band, in agreement with the experiment. This has also been documented in classical molecular dynamics simulations.[42]

As also can be seen in Figure 7, very different amplifications of the intensities of bending (1611 cm^{-1}) and stretching (3069 cm^{-1}) bands can be obtained depending on the type of quantum corrections. The agreement with experiment can either improve or become worse. Over the spectral domain investigated here, the most reasonable infrared spectrum is obtained with application of the $\beta\hbar\omega/(1 - \exp(-\beta\hbar\omega))$ detailed balance factor. This has also been observed by Ahlborn et al.[96] in classical simulations of water. In particular, when multiplied by this correction factor, amplitudes of the calculated infrared bands at 639 and 1611 cm^{-1} become comparable with experiment (Figure 7). The intramolecular bending band shape is in very good agreement, whereas the librational band shape is a bit too narrow in comparison to the measured profile. The intensity of the O−H stretching absorption band is overestimated by $\sim$20%. Still, the $\beta\hbar\omega/(1 - \exp(-\beta\hbar\omega))$ quantum correction is the one coming closest to experiment.

Mode amplitude corrections (accounting for the quantum nature of oscillators) give results equivalent to the effect of the $2/(1 + \exp(-\beta\hbar\omega))$ desymmetrization procedure below 1000 cm^{-1} and are intermediate between the $2/(1 + \exp(-\beta\hbar\omega))$ and $\beta\hbar\omega/(1 - \exp(-\beta\hbar\omega))$ corrections above 1000 cm^{-1}. In particular, this correction is not sufficient to recover the amplitude of the O−H stretching infrared band (3069 cm^{-1}). When desymmetrization and mode amplitude corrections are combined together (Figure 7), the best result is obtained with the $2/(1 + \exp(-\beta\hbar\omega))$ factor. As expected (see section 3), the calculated infrared spectrum of water is then identical to the one obtained with the $\beta\hbar\omega/(1 - \exp(-\beta\hbar\omega))$ desymmetrization procedure alone. When the latter desymmetrization procedure is supplemented with mode amplitude correction, all amplitudes of the infrared spectrum are drastically overestimated with respect to the experiment. We note that under the thermal conditions considered here the $\exp(\beta\hbar\omega/2)$ correction shows an exponential growth that overwhelms everything else. Application of this correction in the spectral domain investigated here

Feature Article

J. Phys. Chem. B, Vol. 107, No. 38, 2003 **10357**

overestimates all amplitudes by a fair amount and thus should be discarded.

Uracil in Vacuo. We have also investigated the effect of quantum corrections on the infrared spectrum of isolated uracil discussed in section 4. For an order of magnitude estimate of the effect of quantum corrections on this system, we mention that multiplication by $2/(1 + \exp(-\beta\hbar\omega))$ only results in a doubling of intensities. Multiplication by the mode amplitude factor gives rise to absolute intensities intermediate between the $2/(1 + \exp(-\beta\hbar\omega))$ and $\beta\hbar\omega/(1 - \exp(-\beta\hbar\omega))$ corrections. Because the units for intensity in the spectrum of ref 83 are not given, we cannot check our absolute values with the experiment and only compare the ratio of amplitudes associated with different peaks of the spectrum. Hence, we found that relative amplitudes of the three main peaks observed in the 2000−1000 cm^{-1} spectral region, that is, at 1619, 1280, and 1068 cm^{-1}, are in good agreement with experiment for the case of the $\beta\hbar\omega/(1 - \exp(-\beta\hbar\omega))$ correction. For example, the ratio between the amplitude of the 1619 and 1068 cm^{-1} bands, which was found to be too small in the uncorrected spectrum (7.97), increases with the application of this quantum correction and becomes 11.91. This is in very good agreement with the 11.33 ratio in the experiment. Moreover, the ratio between the amplitudes of the 1280 and 1068 cm^{-1} bands is reduced, going from 2.50 in the uncorrected spectrum to 2.10 after the correction application. This value is to be compared with the 1.83 in the experimental spectrum. Again, the $2/(1 + \exp(-\beta\hbar\omega))$ factor coupled with the mode amplitude correction $\beta\hbar\omega/(2 \tanh(\beta\hbar\omega/2))$, is equal to $\beta\hbar\omega/(1 - \exp(-\beta\hbar\omega))$ applied by itself. However, when $\beta\hbar\omega/(1 - \exp(-\beta\hbar\omega))$ is supplemented by the mode amplitude correction, an overestimation of the experimental intensities by a factor of ∼1.5 is obtained in our calculation. We conclude that the most satisfactory agreement with experiment in the 2000−1000 cm^{-1} range can be achieved using $\beta\hbar\omega/(1 - \exp(-\beta\hbar\omega))$.

Aqueous Uracil. Finally, we have tested the effect of quantum corrections on the amplitudes of the infrared modes of aqueous uracil. As in the case of isolated molecule, the comparison is between ratios of amplitudes of different vibrational bands, as obtained by our calculations and experiment. Conclusions are not as clear as those for isolated uracil. As in the vacuum case, the mode amplitude correction gives rise to absolute intensities intermediate between the $2/(1 + \exp(-\beta\hbar\omega))$ and $\beta\hbar\omega/(1 - \exp(-\beta\hbar\omega))$ corrections. Relative amplitudes of the 1619 cm^{-1} band and the 1161 cm^{-1} band (i.e., the C=O stretch versus C−H bending band) are overestimated. Experiment gives a value of ∼4; classical calculations yield 4.2. We obtain 5.7, whatever quantum correction is applied. On the other hand, the ratio of the amplitudes of the 1619 and 1441 cm^{-1} bands varies between 7.7 (for the bare value or with the $2/(1 + \exp(-\beta\hbar\omega))$ correction) to 8.1 (with the mode amplitude correction) or 8.6 (using $\beta\hbar\omega/(1 - \exp(-\beta\hbar\omega))$). This is to be compared to ∼8.4 in the experimental spectrum. Combining detailed balance and mode amplitude correction leads to overestimation. The conclusion, therefore, is that if we are willing to make a rather adhoc choice of correction procedure, overall agreement with experiment can be improved, although not by very much.

References and Notes

(1) Car, R.; Parrinello, M. *Phys. Rev. Lett.* **1985**, *55*, 2471.

(2) Marx, D.; Hutter, J. Ab initio molecular dynamics: theory and implementation. In *Modern methods and algorithms of quantum chemistry*; Grotendorst, J., Ed.; NIC Series, Vol. 1; John von Meumann Institute for Computing: Julich, Germany, 2000; pp 301−449.

(3) Frenkel, D.; Smit, B. *Understanding molecular simulation*, 2nd ed.; Academic Press: San Diego, CA, 2002.

(4) Sprik, M.; Hutter, J.; Parrinello, M. *J. Chem. Phys*. **1996**, *105*, 1142.

(5) Silvestrelli, P. L.; Parrinello, M. *J. Chem. Phys*. **1999**, *111*, 3572.

(6) Tuckerman, M.; Laasonen, K.; Sprik, M.; Parrinello, M. *J. Chem. Phys.* **1995**, *103*, 150.

(7) Marx, D.; Tuckerman, M.; Hutter, J.; Parrinello, M. *Nature* **1999**, *397*, 601.

(8) Tuckerman, M. E.; Marx, D.; Parrinello, M. *Nature* **2002**, *417*, 925.

(9) Trout, B. L.; Parrinello, M. *J. Phys. Chem. B* **1999**, *103*, 7340.

(10) Geissler, P., Dellago, C.; Chandler, D.; Hutter, J.; Parrinello, M. *Science* **2001**, *291*, 2121.

(11) Schwegler, E.; Galli, G.; Gygi, F.; Hood, R. Q. *Phys. Rev. Lett.* **2002**, *87*, 265501.

(12) Laasonen, K. E.; Klein, M. L. *J. Phys. Chem. A* **1997**, *101*, 98.

(13) Sillanpaa, A. J.; Simon, C.; Klein, M. L.; Laasonen, K. *J. Phys. Chem. B* **2002**, *106*, 11315.

(14) Meijer, E. J.; Sprik, M. *J. Am. Chem. Soc*., **1998**, *120*, 6345.

(15) Doclo, K.; Roethlisberger, U. *J. Phys. Chem. A* **2000**, *104*, 6464.

(16) Ensing, B.; Meijer, E. J.; Blochl, P. E.; Baerends, E. J. *J. Phys. Chem. A* **2001**, *105*, 3300.

(17) Van Erp, T. S.; Meijer, E. J. *Chem. Phys. Lett.* **2001**, *333*, 290.

(18) Kirchner, B.; Hutter, J. *Chem. Phys. Lett.* **2002**, *364*, 497.

(19) Ramaniah, L. M.; Bernasconi, M.; Parrinello, M. *J. Chem. Phys.* **1998**, *109*, 6839.

(20) Ramaniah, L. M.; Bernasconi, M.; Parrinello, M. *J. Chem. Phys.* **1999**, *111*, 1587.

(21) Vuilleumier, R.; Sprik, M. *J. Chem. Phys.* **2001**, *115*, 3454.

(22) Lyubartsev, A. P.; Laasonen, K.; Laaksonen, A. *J. Chem. Phys.* **2001**, *114*, 3120.

(23) Marx, D.; Sprik, M.; Parrinello, M. *Chem. Phys. Lett.* **1997**, *273*, 360.

(24) Bako, I.; Hutter, J.; Palinkas, G. *J. Chem. Phys.* **2002**, *117*, 9838.

(25) Pasquarello, A.; Petri, I.; Salmon, P. S.; Parisel, O.; Car, R.; Toth, E.; Powell, D. H.; Fischer, H. E.; Helm, L.; Merbach, A. E. *Science* **2001**, *291*, 856.

(26) Ensing, B.; Buda, F.; Bloechl, P. E.; Baerends, E. J. *Phys. Chem. Chem. Phys.* **2002**, *4*, 3619.

(27) Bruge, F.; Bernasconi, M.; Parrinello, M. *J. Am. Chem. Soc.* **1999**, *121*, 10833.

(28) Sillanpaa, A. J.; Paivarinta, J. T.; Hotokka, M. J.; Rosenholm, J. B.; Laasonen, K. E. *J. Phys. Chem. B* **2001**, *105*, 10111.

(29) Tobias, D. J.; Jungwirth, P.; Parrinello, M. *J. Chem. Phys.* **2001**, *114*, 7036.

(30) Meijer, E. J.; Sprik, M. *J. Phys. Chem. A* **1998**, *102*, 2893.

(31) Alber, F.; Folkers, G.; Carloni, P. *J. Phys. Chem. B* **1999**, *103*, 6121.

(32) Kuo, I. F.; Tobias, D. J. *J. Phys. Chem. B* **2001**, *105*, 5827.

(33) Hutter, J.; Carloni, P.; Parrinello, M. *J. Am. Chem. Soc.* **1996**, *118*, 8710.

(34) Molteni, C.; Parrinello, M. *J. Am. Chem. Soc.* **1998**, *120*, 2168.

(35) Carloni, P.; Sprik, M.; Andreoni, W. *J. Phys. Chem. B* **2000**, *104*, 823.

(36) Aktah, D.; Frank, I. *J. Am. Chem. Soc.* **2002**, *124*, 3402.

(37) Gervasio, F. L.; Carloni, P.; Parrinello, M. *Phys. Rev. Lett.* **2002**, *89*, 108012.

(38) Carloni, P.; Rothlisberger U.; Parrinello, M. *Acc. Chem. Res.* **2002**, *35*, 455.

(39) Reimers, J.; Watts, R. O. *Chem. Phys.* **1984**, *91*, 201.

(40) Toukan, K.; Rahman, A. *Phys. Rev. B* **1984**, *31*, 2643.

(41) Madden, P. A.; Impey, R. W. *Chem. Phys. Lett.* **1986**, *123*, 502.

(42) Guillot, B. *J. Chem. Phys.* **1991**, *95*, 1543.

(43) Corongiu, G.; Clementi, E. *J. Chem. Phys.* **1993**, *98*, 4984.

(44) Marti, J., J.; Padro, J. A.; Guardia, E. *J. Chem. Phys.* **1996**, *105*, 639.

(45) Vuilleumier, R.; Borgis, D. *J. Chem. Phys.* **1999**, *111*, 4251.

(46) Chelli, R.; Cardini, G.; Procacci, P.; Righini, S.; Califano, S.; Albrecht, A.; *J. Chem. Phys.* **2000**, *113*, 6851.

(47) Stern, H. A.; Berne, B. J. *J. Chem. Phys.* **2001**, *115*, 7622.

(48) Iuchi, S.; Morita, A.; Kato, S. *J. Phys. Chem. B* **2002**, *106*, 3466.

(49) Silvestrelli, P. L.; Bernasconi, M.; Parrinello, M. *Chem. Phys. Lett.* **1997**, *277*, 478.

(50) Bernasconi, M.; Silvestrelli, P. L.; Parinello, M. *Phys. Rev. Lett.* **1998**, *81*, 1235.

(51) Putrino, A.; Parrinello, M. *Phys. Rev. Lett.* **2002**, *88*, 176401.

(52) Putrino, A.; Sebastiani, D.; Parrinello, M. *J. Chem. Phys.* **2000**, *113*, 7102.

(53) King-Smith, R. D.; Vanderbilt, D. *Phys. Rev. B* **1993**, *47*, 1651.

(54) Vanderbilt, D.; King-Smith, R. D. *Phys. Rev. B* **1993**, *48*, 4442.

(55) Martin, R. *Phys. Rev. B* **1974**, *9*, 1998.

(56) Resta, R. *Rev. Mod. Phys.* **1994**, *66*, 899

(57) Resta, R. *J. Phys.: Condens. Matter* **2000**, *12*, R107.

(58) Resta, R. *Phys. Rev. Lett.* **1998**, *80*, 1800.

(59) Marzari, N.; Vanderbilt, D. *Phys. Rev. B* **1997**, *56*, 12847.

(60) Silvestrelli, P. L.; Parrinello, M. *Phys. Rev. Lett*. **1999**, *82*, 3308.

(61) Souza, I.; Martin, R. M.; Marzari, N.; Zhao, X.; Vanderbilt, D. *Phys. Rev. B* **2000**, *62*, 15505.

(62) Taillandier, E.; Peticolas, W. L.; Adam, S.; Huynh-Dinh, T.; Igolen, J. *Spectrochim. Acta A* **1990**, *46A*, 107.

(63) Taillandier, E.; Liquier, J.; Ghomi, M. *J. Mol. Struct.* **1989**, *214*, 185.

(64) Powell, J. W.; Peticolas, W. L.; Genzel, L. *J. Mol. Struct.* **1991**, *247*, 107.

(65) Fritzche, H.; Pohle, W. *J. Mol. Struct*. **1990**, *219*, 341.

(66) Fritzche, H. *J. Mol. Struct*. **1991**, *242*, 245.

(67) Semenov, M.; Bolbukh, T.; Mallev, V. *J. Mol. Struct*. **1997**, *408−409*, 213.

(68) Ghomi, M.; Aamouche, A.; Cadioli, B.; Berthier, G.; Grajcar, L.; Baron, M. H. *J. Mol. Struct*. **1997**, *411*, 323.

(69) Aamouche, A.; Berthier, G.; Cadioli, B.; Gallinella, E.; Ghomi, M. *J. Mol. Struct. (THEOCHEM)* **1998**, *426*, 307.

(70) Aamouche, A.; Ghomi, M.; Coulombeau, C.; Jobic, H.; Grajcar, L.; Baron, M. H.; Baumruk, V.; Turpin, P. Y.; Henriet, C.; Berthier, G. *J. Phys. Chem*. **1996**, *100*, 5224.

(71) Gaigeot, M.-P.; Sprik, M. *J. Phys. Chem. B*, submitted for publication.

(72) McQuarrie, D. A. *Statistical Mechanics*; Harper-Collins Publishers: New York, 1976.

(73) Resta, R.; Sorella, S. *Phys. Rev. Lett* **1999**, *82*, 370.

(74) Souza, I.; Marzari, N.; Vanderbilt, D. *Phys. Rev. B* **2001**, *65*, 035109.

(75) Boys, S. F. In *Quantum Theory of Atoms, Molecules, and the Solid State*; Löwdin, P. O., Ed.; Academic Press: New York, 1966; p 253.

(76) Silvestrelli, P. L.; Marzari, N.; Vanderbilt, D.; Parrinello, M. *Solid State Commun*. **1998**, *107*, 7.

(77) Silvestrelli, P. L. *Phys. Rev. B* **1999**, *59*, 9703.

(78) Berghold, G.; Mundy, C. J.; Romero, A. H.; Hutter, J.; Parrinello, M.; *Phys. Rev. B* **2000**, *61*, 10040.

(79) Trouillier, N.; Martins, J. L. *Phys. Rev. B* **1991**, *43*, 1993.

(80) Kleinman, L.; Bylander, D. M. *Phys. Rev. Lett*. **1982**, *48*, 1425.

(81) Becke, A. *Phys. Rev. A* **1988**, *38*, 3098.

(82) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.

(83) Les, A.; Adamowicz, L.; Nowak, M. J.; Lapinski, L. *Spectrochim. Acta* **1992**, *48A*, 1385.

(84) Hutter, J.; Alavi, A.; Deutsch, T.; Bernasconi, M.; Goedecker, S.; Marx, D.; Tuckerman, M.; Parinello, M. *CPMD*, version 3.4.1; IBM Research Division, MPI Festkoerperforschung: Stuttgart, Germany.

(85) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*, revision A.7; Gaussian, Inc.: Pittsburgh, PA, 1998.

(86) Chandra, A. K.; Nguyen, M. T.; Zeegers-Huyskens, T. *J. Phys. Chem. A* **1998**, *102*, 6010.

(87) VanMourik, T.; Price, S. L.; Clary, D. C. *J. Phys. Chem. A* **1999**, *103*, 1611.

(88) Gaigeot, M. P.; Ghomi, M. *J. Phys. Chem. B* **2001**, *105*, 5007.

(89) Gaigeot, M. P.; Kadri, C.; Ghomi, M. *J. Mol. Struct*. **2001**, *565−566*, 469.

(90) Stewart, R. F.; Jensen, L. H. *Acta Crystallogr*. **1967**, *23*, 1102.

(91) Ferenczy, G.; Harsany, L.; Rozsondai, B.; Hargittai, I. *J. Mol. Struct*. **1986**, *140*, 71.

(92) Brown, R. D.; Godfrey, P. D.; McNaughton, D.; Pierlot, A. P. *J. Am. Chem. Soc*. **1988**, *110*, 2329.

(93) Rusk, A. N.; Williams, D.; Querry, M. R. *J. Opt. Soc. Am*. **1971**, *61*, 895.

(94) Afsar, M. N.; Hasted, J. B. *J. Opt. Soc. Am*. **1977**, *67*, 902.

(95) Borysow, J.; Moraldi, M.; Frommhold, L. *Mol. Phys*. **1985**, *56*, 913.

(96) Ahlborn, H.; Space, B.; Moore, P. B. *J. Chem. Phys*. **2000**, *112*, 8083.