How to Model Solvation of Peptides? Insights from a Quantum-mechanical and Molecular Dynamics Study of *N*-Methylacetamide. 1. Geometries, Infrared, and Ultraviolet Spectra in Water

Benedetta Mennucci*,† and José M. Martínez‡

Dipartimento di Chimica e Chimica Industriale, Università di Pisa, via Risorgimento 35, 56126 Pisa, Italy, and Departamento de Química Física, Universidad de Sevilla, Facultad de Química, Sevilla 41012, Spain

Received: January 4, 2005; In Final Form: March 4, 2005

This paper represents the first part of a study of solvation in peptides using quantum-mechanical and classical approaches. In this study, the peptide is modeled as its simplest analogue, namely, *N*-methyl-acetamide, and the effects of the solvent (here water, and in the second part of the study, water and acetone) are introduced at three different levels, e.g., through a continuum description, using solute−solvent clusters, and using the same clusters embedded in an external continuum. In turn, the solute−solvent clusters have been obtained in two alternative ways, either by using QM optimization procedures or extracting a proper set of structures from MD simulations. In this part of the study, geometries, IR, and UV spectra are calculated in terms of the different solvation models, and the results are analyzed and compared to get insigths about different aspects of solvation involving dynamic and static effects on one hand and bulk or specific interactions on the other hand.

1. Introduction

The study of solvation of proteins and peptides is an important issue in molecular biology and biochemistry. Contrary to numerous experimental reports on the solvation of peptides and proteins (both in crystals and in liquid solutions)[1−6] and the large variety of theoretical studies on solvent-induced modifications of properties and reactivity of biomolecular systems (see ref 7 for a review on the argument and for a selection of literature), accurate quantum-mechanical (QM) studies on the nature of solvation of proteins are rather limited. This kind of study is in fact extremely complex but such complexity can be significantly reduced by utilizing the group concept. For proteins the peptide group is the repetitive moiety, and thus, it is natural to start with it to analyze the nature and the effects of solvation in proteins.

The simplest model for the amide linkage of peptides (and proteins) is *N*-methyl-acetamide (NMA). The interest of experimentalists on such a simple system is double. First a clear understanding of the NMA structure is considered as the basis for understanding the geometric constraints imposed by the peptide linkages that determine, at least partly, the protein structure. Second, a detailed understanding of NMA spectroscopic features is assumed as the fundamental basis for spectroscopic methods to monitor protein structure and dynamics. Both of these expectations are of real interest and of future applications only if both structures and spectroscopic properties of NMA are monitored in the presence of the natural environment of the biological systems. In particular, it is of fundamental importance the study of hydration and of hydrogen bonding effects.

From the theoretical point of view, NMA represents an ideal model as small enough to be treated in a very accurate way through QM approaches.[8−23] Conversely, the study of its solvation at the same level of accuracy is still an impossible task, and thus approximations have to be introduced. Two ways can be followed for that purpose: either we introduce approximations in the description of intra- and intermolecular interactions, or we reduce the number of molecules to be treated. In the first case, the most suitable approaches are represented by classical simulations (mainly Monte Carlo and molecular dynamics, MD), whereas in the second case the so-called supermolecule approach is generally exploited. The latter approach extends the accurate (usually QM) description besides the solute by including a reduced number of solvent molecules, namely those which should represent the first solvation shell(s). An alternative approach to these two possibilities is represented by solvation continuum models.[24,25] The potentiality of this approach, in which the solvent loses its molecular nature which is replaced by a structureless continuum description, is due not only to the fact that the dimension and in the complexity of the problem to be treated remain almost unchanged with respect to the gas phase but also to its ability to properly account for both polarization and long-range effects. These two effects cannot (at least not easily) be simultaneously taken into full account by classical simulations or supermolecule approaches. In addition, modern continuum solvation models can be easily extended to include high-level QM descriptions and also generalized to compute various quantities such as equilibrium geometries and response properties with respect to applied electric and magnetic fields.[26] Continuum models, however, cannot completely describe an important part of solvation, namely that related to local and/or specific interactions such as, for example, solute−solvent hydrogen bonding. To overcome such a limitation, continuum models can be supplemented with specific correction terms,[27] or more commonly, coupled to supermolecule descriptions and/or to classical simulations.

These studies are useful not only to obtain accurate descriptions of solvated systems and of their properties but also to

* To whom correspondence should be addressed. E-mail: bene@dcci.unipi.it.
† Università di Pisa.
‡ Universidad de Sevilla.

10.1021/jp050034z CCC: $30.25
Published on Web 04/12/2005

investigate methodological and interpretative aspects. In fact, the analysis of limits and potentialities of different solvation models represents a necessary step to clarify and interpret different aspects of solvation. This kind of analysis, however, requires that we can identify some experimentally well-determined quantities that can be easily related to solvent effects and that, at the same time, can be reliably reproduced by theoretical models. These quantities are represented by electronic, vibrational, and magnetic spectroscopies of molecular solutes in dilute solutions.[28−30]

In the two papers we devote to the solvation of NMA, experimental data on IR, UV, and NMR spectroscopies of NMA in the gas phase, in water, and in acetone will be used as reference values to compare with different solvation approaches, namely continuum-only, QM supermolecule in the gas phase, QM supermolecule coupled to the continuum, and supermolecule obtained through a statistical average on configurations obtained from MD simulations (once again in the gas phase and coupled to the continuum). The relative performances of the different models will be then used for an analysis on the nature of NMA solvation. In all cases, the continuum solvation model is represented by the integral equation formalism (IEF)[31] version of the polarizable continuum model (PCM).[32]

More specifically, the present paper (from now on indicated as paper 1) will be focused on the effects of hydration on structures, IR spectra, and UV absorptions of NMA, whereas in the second paper (paper 2),[33] the analysis will be shifted to solvation effects on ^{17}O and ^{15}N nuclear shieldings, and besides water, acetone will also be considered as solvent. It is worth pointing out that the main aim of these works is not only the reproduction of the experimental data, but instead the use of these spectroscopic probes to asses the limitations and potentialities of the different solvation models here considered in order to shed light on both the physics and the dynamics of the NMA solvation in protic and nonprotic polar solvents. The methodology employed to achieve that goal allows the inclusion of different contributions in a stepwise way.

2. Computational Details

2.1. Quantum-Mechanical Calculations. *Geometry Optimizations.* For all of the systems (single NMA and NMA−water clusters) both in vacuo and in solution, geometry optimizations were performed at the density functional theory (DFT) using the hybrid functional which mixes the Lee, Yang, and Parr functional for the correlation part and Becke's three-parameter functional for the exchange (B3LYP).[34] The basis set used was the 6-31+G(d,p).

UV Spectra. To evaluate absorption energies, the time-dependent density functional theory (TDDFT) was used with the B3LYP functional and the Dunning basis set (d95 according to Gaussian notation). The calculations on NMA and on QM NMA-nw structures were performed adding one diffuse (+) and one polarization (d) function on all O, N, and C atoms, whereas those on MD structures were obtained exploiting a mixed basis set: d95v+(d) for the NMA molecule and d95v for the solvent molecules (this scheme will be indicated as hybrid).

Solvation. In the IEF−PCM solvation model, the solvent electrostatic reaction field is represented in terms of an apparent surface charge defined on the surface of the molecular cavity (i.e., the boundary between solute and solvent). In the IEF−PCM model, the molecular cavity is obtained in terms of interlocking spheres centered on selected nuclei. The chosen radii are 1.7 for carbonyl carbon, 2.0 for the methyl group, 1.6 for N, 1.52 for O, and 1.2 for H of H-bonded water.[35] All of the radii have been multiplied by a factor (equal to 1.2 if not otherwise specified) in order to take into account the fact that atomic bond or lone pair centers of the solvent molecules are normally located a bit further from the solute atoms than a van der Waals radius.[24]

In general, the solute electronic and nuclear charge distribution and solvent reaction field are allowed to mutually equilibrate; however, when the solute abruptly changes its electronic state through a vertical transition, the relaxation of the reaction field in the direction of the new solute state may be incomplete. By taking into account the typical time scales characterizing electronic and nuclear (or molecular) motions, we assume that only the part of the solvent reaction which is induced by the polarization of its electrons can immediately modify according to the new electronic state reached by the solute in the transition process; all of the rest remains frozen in the previous equilibrium condition determined by the initial state. In a reasonable approximation, the fast component can be taken to be proportional to the dielectric constant at infinite frequency ϵ_∞ where $\epsilon_\infty \approx n^2$ and n is the refractive index of the solvent. In the framework of IEF−PCM, this scheme is implemented introducing two sets of apparent charges representing the electronic (or fast) and the slow contributions of the solvent reaction, respectively.[36] IEF−PCM has been recently generalized to TDDFT calculations also to include solvent perturbation operators in the coupled perturbed scheme; for more details on the formalism see ref 37. The values used in the following calculations for ϵ and ϵ_∞ of water are 78.39 and 1.776, respectively.

All QM calculations both in vacuo and in solution were performed using a local version of the Gaussian code.[38]

2.2. MD Simulations. *Simulation Details.* MD simulations were performed with the DL_POLY 2.13 package[39] in the NVE ensemble using periodic boundary conditions. The flexible all-atom AMBER94 force field[40] was used to model NMA molecule as implemented in ref 41, whereas the TIP3P[42] model was adopted for water. Simulations with rigid and flexible NMA geometry were performed. In the case of the simulation with rigid geometries, the NMA molecule was treated as a rigid body and only intermolecular interactions were considered. In this case, geometry IV optimized in water was adopted as the solute geometry (see section 3). Systems were thermalized at 300 K for 30 ps, and a 300 ps production period was used for analysis. Simulations contained one NMA molecule and 399 solvent molecules in a cubic box reproducing the experimental water density (0.997 g/cm^3). Equations of motions were integrated with time steps of 0.2 and 2 fs for the flexible and rigid cases, respectively. Coulombic interactions were computed using the smoothed particle mesh Ewald method,[43] whereas for the short-range nonbonded interactions, a cutoff of $L/2$ was applied.

Radial Distribution Functions and Selection of Structures. As described in section 1, the information obtained from the MD analysis will be used to define solute−solvent clusters to be studied through quantum mechanics computations. In particular, using radial distribution functions (RDFs), cutoff distances are defined from selected solute atoms to solvent molecules to obtain clusters including the desired number of solvation shells.[28] As the interest is here focused on H-bonding effects, the RDFs we will consider will be those centered on carbonyl oxygen and amide hydrogen atoms of NMA. The remaining two methyl groups are involved in much weaker C−H···OW (W = water) interactions[44] which will not be considered in the present study.

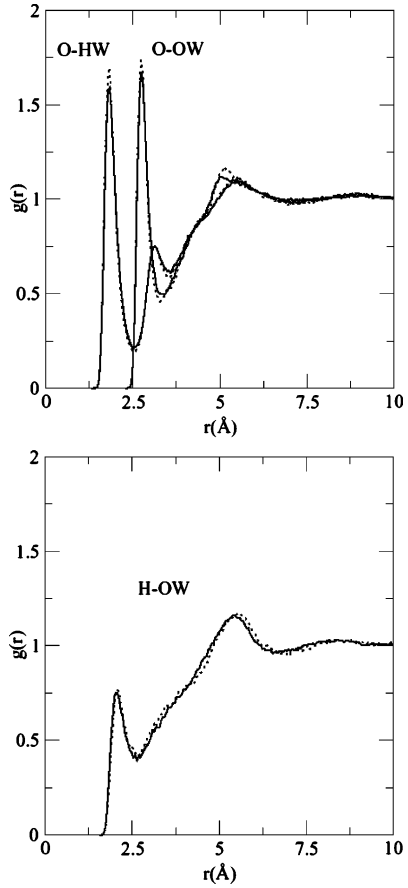

Figure 1. (C)O−HW,(C)O−OW and (N)H−OW RDFs obtained from the two simulations performed on NMA in water, namely, that containing a fully flexible solute (full line) and that in which the NMA structure is kept fixed at the QM geometry (conformer IV) (dotted line).

Figure 1 shows the (C)O−HW, (C)O−OW, and (N)H−OW RDFs obtained from the two simulations performed of NMA in water, namely, that containing a fully flexible solute and that in which the NMA structure is fixed to the QM geometry (conformer IV, see section 3).

By comparing full and dotted lines, we can observe that freezing the geometry of NMA has only a marginal effect on the hydration structure around the carbonyl and amide groups. The O−HW RDF shows a very well-defined first peak centered at 1.83 Å that integrates to 2.3 (2.73 Å and 2.6 for the O−OW distribution). Therefore, on average, two hydrogen atoms of two different water molecules hydrate the carbonyl group by means of hydrogen bonding. The (N)H−OW distribution defines a much less structured solvent around the amide group. The first peak, centered at 2.05 Å is not so sharply defined and the running integration number up to the first minimum is 1.1. These results, in agreement with previous MD simulations[45−50] using a wide range of models, allow us to select the clusters to be used in QM calculations: water molecules having any H atom closer to (C)O atom than the corresponding first minimum of the O−HW RDF will be included in the cluster. An analogous criterion is applied for the (N)H−OW pair. The selected structures (see section 5) were sampled at equal time intervals long enough (above 5 ps) as to avoid correlation, so a proper sampling can be performed on the basis of the configurations used.

3. Tautomers: Structures and Energies

By means of IR and NMR spectroscopies, it appears that both in the gas phase and in solution, the cis and trans isomers arising from the rotation around the CN bond are not equally abundant but instead *trans*-NMA is dominant.[51,52] Following these experimental data, here we shall limit the study to the trans form only. For this isomer the outcome of the stationary point search and vibrational analysis at the B3LYP/6-31+G(d,p) level consisted of four conformers characterized by the value of the torsional angles φ and ψ. The structures are reported in Figure 2.

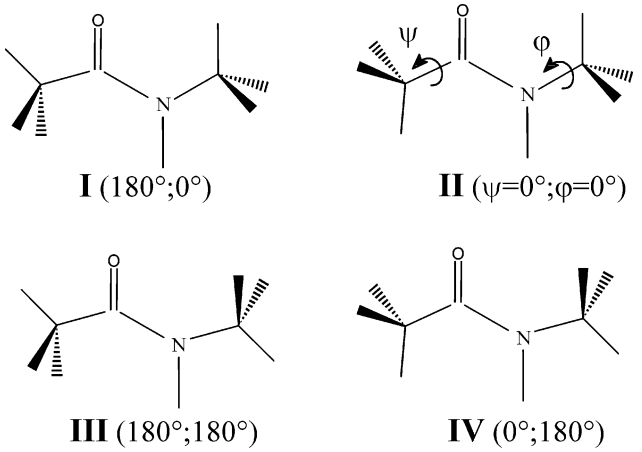

Figure 2. NMA conformers obtained by rotating the two methyl groups according to the angles ψ and φ.

TABLE 1: $\Delta G = G_X - G_{min}$ (kcal/mol) and normalized Boltzmann Weights (in Parentheses) of the Four Conformers (X = I, II, III, and IV) in the Gas Phase (vac) and in Water[a]

	vac	water
I	0.98 (0.139)	1.87 (0.029)
II	0.00 (0.728)	0.87 (0.155)
III	2.06 (0.022)	0.92 (0.142)
IV	1.12 (0.111)	0.00 (0.675)

[a] Zero of energy is taken with respect to the most stable form in each phase.

In Table 1, we report the relative thermal free energies of the four structures (with respect to the most stable one in each phase) and the corresponding normalized Boltzmann weights.

As shown by results reported in Table 1, the relative stabilities of the four conformers largely depends on the solvation; although structure II is the most stable in the gas phase, structure IV becomes the most stable in water. The vibrational analysis of these structures reveals that only structures II and IV are real minima (they present all positive eigenvalues) in the gas phase and in solution, respectively; all of the other structures present imaginary frequencies corresponding to methyl rotations. Previous QM calculations[8,10,54] have shown that in the gas phase the lowest energy structure differs for different levels of calculation and for different basis sets. To properly check this aspect, we have repeated geometry optimizations and frequency calculations at B3LYP/6-311+G(d,p) and MP2/6-311+G(d,p) levels for the four conformers of isolated NMA. The effect of increasing the basis set leads, as expected, to a shift of structure IV as the second most stable structure immediately after structure II which still remains the lowest-energy conformer (and the only real minimum). Passing to MP2, the situation reverts: structure IV becomes the lowest (and real) stationary state and structure II is the second most stable conformer (but with a negative frequency); in this case, ΔG between IV and II conformers is 0.48 kcal/mol. All of these results seem to indicate that structures II and IV are basically equivalent minimum-energy states for isolated NMA (we note that experiments seem to indicate an almost free rotation around the angle φ at room

TABLE 2: Comparison of Calculated (B3LYP/6-31+G(d,p)) Geometries of NMA in the Gas Phase (Conformer II) and in Water (Conformer IV) with Experiments[a]

	calc		calc	
	vac	water	electron diffraction[b]	crystal[c]
CO	1.229	1.243	1.225	1.236
CN	1.368	1.351	1.386	1.290
NH	1.009	1.010	(1.002)	
$NC_{m'}$	1.454	1.458	1.469	1.465
CC_m	1.520	1.515	1.520	1.536
$\angle C_mCN$	115.6	116.9	114.1	116.5
$\angle OCN$	122.8	121.7	121.8	123.0
$\angle CNH$	118.4	118.9	110.0	119.0
$\angle CNC_{m'}$	123.0	122.4	119.7	120.5
$\psi(HC_mCN)$	0.4	1.3		0
$\varphi(CNC_{m'}H)$	0.4	179.5		180.0
OCNH	180.0	180.0	180.0	180.0
$C_mCNC\ m'$	180.0	180.0	180.0	180.0

[a] Bond lengths are in angstrom, and angles are in degrees. The indices m and m' distinguish between methyl carbon bonded to C(O) and methyl carbon bonded to N(H). [b] From refs 58 and 85. [c] From refs 56 and 57.

temperature[55]). For hydrated NMA, any level of calculations (either B3LYP or MP2, and various basis sets) indicates conformer IV as the minimum energy state (and the real minimum), a fact confirmed by previous calculations[10,54] and by crystallographic data.[56,57]

In Table 2, we report a selection of geometrical parameters for the minimum structures of NMA in the gas phase and in water together with available experimental data. The two sets of experimental data refer to electron diffraction and X-ray measurements, in the gas phase[58] and on crystals,[56,57] respectively. Previous calculations[10,18,59] have shown that the equilibrium structure of isolated trans NMA has a planar symmetry: this is confirmed also by the present calculations as shown in Table 2.

The data for the isolated system (VAC) when compared to the experimental electron diffraction values show a good correlation which make us confident with the B3LYP/6-31+G(d,p) level of calculation.

As already indicated, the $\varphi(CNC_{m'}H)$ dihedral angle passes from 0° in the gas phase to ~180° in solution, this can be related to the more favorable intramolecular interaction between the lone pair electrons of the oxygen and the eclipsed hydrogen of the methyl group in the isolated NMA, reduced in solution by interactions with the solvent. Besides the rotation of the methyl group, effects of solvation are significant also in other parameters. In particular, we observe an important increase in the CO bond length passing from gas to water; this trend correctly reproduces the gas-to-crystal experimental change as well as the red-shift observed in the related stretching frequency in the IR spectrum of the water solution (a detailed analysis on the IR spectra in gas and in water solution will be reported in the following section). Another important change is the parallel decrease of CN bond.

4. Continuum Description

4.1. IR Frequencies. Although many experimental and theoretical works have been performed on the vibrational analysis of NMA,[8,17,51,54,59−71] some uncertainties and even contradictions remain. Here we shall focus on the changes in the frequencies of four fundamental bands (the so-called amide I, amide II, amide III, and amide NH bands) when passing from the gas phase to a water solution at infinite dilution.

Amide I (AI) generally appears as the most intense band in the IR spectra of peptides and proteins.[72] It is usually described as a pure CO stretching mode, even if the normal-mode pattern is more complex, especially for NMA in water. Chen and coworkers[63] provided experimental and theoretical evidences for vibrational mixing (giving rise to two broad subbands) between AI and the bending modes of the hydrogen bonded water molecules. In addition, the influence of the aqueous solvent is important on the AI wavenumber, which is downshifted from a the gas phase value of ca. 1720 cm^{-1} (1728 according to Mayne and Hudson[61] and 1723 Kubelka and Keiderling,[17]) to ca. 1630 cm^{-1}.[66] Amide II is described as an out-of-phase combination of CN stretching and CNH bending. In the gas phase, this band is observed[61] at 1500 cm^{-1}, whereas in water it is shifted[17] to 1582 cm^{-1}. Amide III (AIII) is generally described as the counterpart to amide II, being primarily CNH bending plus CN stretching. In the spectrum[51] of aqueous NMA, it appears as a single band at 1316 cm^{-1}, whereas in vacuo, it is observed at ca. 1260 cm^{-1} (1265 cm^{-1} according to ref 51 and 1257 cm^{-1} according to ref 61). The NH stretching mode has a very complicated profile, and it is highly temperature sensitive. The form and position of the corresponding band is also largely affected by hydrogen bonding. Experimental data show a large red shift (~160 cm^{-1}) in the frequency passing from aprotic and apolar solvents to water.[66]

TABLE 3: B3LYP/6-31+G(d,p) and Experimental Frequencies (cm^{-1}) of the Four Fundamental Bands of NMA in the Gas Phase (vac) and in Water[a]

	vac	water	shift	vac	water	shift
	AI (CO)			AII		
I	1757	1675		1567	1579	
II	**1751**	1669		**1566**	1579	
III	1754	1674		1549	1560	
IV	1748	**1666**	−85	1549	**1561**	−5
exp[b]	1723	1625	−98	1500	1582/1575[c]	+82/+75
	AIII			N−H		
I	**1275**	1298		**3642**	3654	
II	1273	1296		3653	3664	
III	1283	1304		3655	3667	
IV	1282	**1301**	+28	3665	**3676**	+23
exp[b]	1265	1316	+51	3476/3460[d]	3314[e]	−162/−146

[a] The calculated shifts are with respect to the frequency values of the most stable conformer (in bold) in each phase. [b] From ref 17. [c] Average value of two peaks at 1584 and 1566 cm^{-1}.[64] [d] In dilute CCl4 (left) and in dilute CH_2Cl_2 (right).[66] [e] Raman data in dilute water.[66]

The B3LYP/6-31+G(d,p) results obtained for the four conformers I−IV in the gas phase and in water (i.e., including the IEF continuum) are shown in Table 3. Frequency values for each conformer are reported together with the gas-to-water shifts using the frequencies of the two most stable structures as reference values in the gas phase and with the IEF continuum (namely structure II and IV which are in bold in the table).

It has to be recalled that, due to the neglect of anharmonicity effects and to the approximate treatment of electron correlation, QM vibrational frequencies are typically larger than those observed experimentally: optimum frequency scale factors have been determined for many QM levels of theory.[73] In particular for B3LYP/6-31G(d), a factor of 0.9614 has been suggested;[73c] here however we have preferred to report unscaled frequencies not only because the factor for B3LYP/6-31+G(d,p) is not available but also because these factors have been obtained through fitting procedures on gas-phase data and thus it is not guaranteed that are still valid for solvated systems.[74] The problem of the scaling has however to be kept in mind in the following analysis.

Among the four modes, CO stretching is the one which is less affected by rotational effects: the corresponding frequencies are almost identical in all conformations. All of the other modes, on the contrary, show significant dependencies on the rotations of the two methyl groups. In particular, AII and AIII modes present a clear dual behavior, one for the I,II conformers and the other for III,IV conformers: this means that the rotation affecting the mode is that defined by $\varphi(CNC_{m'}H)$.

The observed frequency shifts when including the effect of the IEF continuum are very different for the different modes. For AI, the computed gas-to-water shift not only presents the correct sign but it is also quite close to the experimental value. This result was not unexpected: from the geometrical parameters reported in Table 2, it clearly appeared that effects of the continuum lead to a longer CO bond and thus, as a consequence, to a lower CO stretching frequency.

Also, the computed shift for AIII presents the right sign but this time the quantitative agreement with experiments is by far worse, as the IEF continuum model significantly underestimates the observed shift. For both AII and NH modes, IEF calculations give a clearly wrong picture (either giving an almost null shift and a wrong sign shift). It is easy to predict that these problems are strongly related to the absence in the IEF model of local and specific interactions between NMA and water molecules and especially hydrogen bonding effects. These in fact when acting on the (N)H can not only modify the NH stretching mode but also the AII mode both through a direct effect on the CNH bending which is part of this mode but also indirectly by modifying the CN distance.

4.2. Absorption Energies. The electronic absorption spectrum of NMA has been measured in the gas phase[75] and in a variety of solvents.[76] The gas-phase spectrum presents two well-separated bands from 6.5 to 7.8 eV (one broad band centered around 6.8 eV and one sharp peak at 7.7 eV). These bands have been assigned to transitions in the π system of the amide group and will be referred to hereafter as the π,π^* transitions. In nonpolar solvents such as cyclohexane or dioxane, a n,π^* also appears as a shoulder. The lowest π,π^{*q} transition comprises an electron transfer from the nitrogen to the carbonyl carbon atom. This transition is red shifted from the gas phase to polar solution and to water. However, in contrast to other amides where large red shifts of up to ~0.5 eV are found, in NMA, the π,π^* transition undergoes a small red-shift (0.1 eV). The n,π^* transition is not measured in the gas phase, and so, the corresponding experimental solvatochromic shift of this transition is not known. However, a blue shift of 0.2 eV is observed passing from cyclohexane to water. Obviously, we cannot assume that the results obtained in cyclohexane are the same as those in the gas phase, for example recent experimental work has shown that NMA tends to self-associate in apolar solvents.[77,78] In any case, we can expect that passing from gas to water a blue shift will be observed. In fact, the n,π^* transition reduces the permanent dipole moment of the corresponding excited state which is thus less strongly stabilized in polar solvents with respect to the ground state.

In Figure 3, orbitals involved in the two transitions as obtained at the B3LYP/d95v+(d) level are plotted. We note that the character of the transitions does not significantly change passing from isolated to (IEF) solvated NMA; in both cases, in fact, the two transitions contain as main contribution a transition from the HOMO for the n,π^* and HOMO-1 for the π,π^* to the same final virtual orbital LUMO+x where x is equal to 2 in the gas phase and equal to 1 in the presence of the IEF continuum.

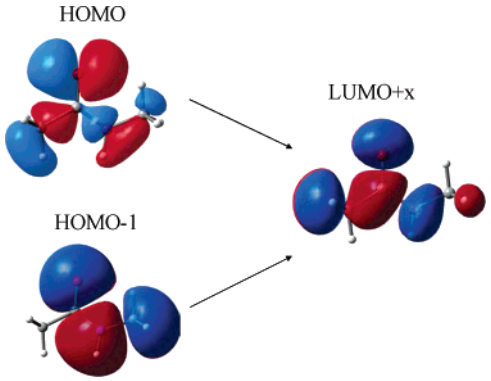

Figure 3. Plot of the orbitals involved in the two transitions as obtained at B3LYP/d95v+(d) level: the two transitions contain, as main contribution, an excitation from the HOMO for the n,π^* , and HOMO-1 for the π,π^*, to the same final virtual orbital LUMO+x where x is equal to 2 in the gas phase and equal to 1 in the presence of the IEF continuum.

TABLE 4: Absorption Energies (eV) and Oscillatory Strengths Computed at d95v+(d) of the Four Conformers in the Gas Phase (vac) and in Water[a]

	$n \rightarrow \pi^*$		$\pi \rightarrow \pi^*$			
	vac	water	vac		water	
	energy	energy	energy	f	energy	f
I	5.62	5.79	7.08	0.0717	6.92	0.2459
II	**5.67**	5.83	**7.12**	0.1178	7.00	0.2726
III	5.55	5.74	7.06	0.0805	6.92	0.2347
IV	5.61	**5.77**	7.13	0.1196	**7.02**	0.2542
δ^{vac}		0.10			−0.10	
δ^{vac}_{exp}		0.2 (cyc)			−0.14	

[a] Values in bold refer to the most stable conformers and the shift refer to them. Exp: $n\pi^*$: 5.5 (eV) (in cyclohexane), 5.7 (eV) (in water); $\pi\pi^*$: vac 6.81 eV, water 6.67 eV;

In Table 4, we report the absorption energies and the oscillatory strengths computed at TDDFT:B3LYP/d95v+(d) level for the four conformers in the gas phase and in water together with the resulting gas-to-water shifts.

Repeating the same scheme of analysis followed for IR frequencies, we first note that the effects of methyl rotations are not completely negligible. In particular, for the $\pi \rightarrow \pi^*$ transition, two groups of similar systems can be identified, namely I,III and II,IV, which seems to indicate that this transition is affected by the rotation around $\psi(HC_mCN)$. This behavior can be better understood looking at the orbital plots reported in Figure 3: the HOMO-1 does not involves the methyl (C_m) moiety but the LUMO+x orbital is significantly spread on it and thus its rotation will affect the relative energy of this orbital and thus the transition energy. No straightforward behaviors can be found for the effects of rotation on $n \rightarrow \pi^*$ transition as in these case, the involved orbitals present large contributions from both methyl groups.

When solvent effects are included, we first note that for both transitions, IEF results reproduce the correct trend, namely a blue-shift on $n \rightarrow \pi^*$ and a red-shift on $\pi \rightarrow \pi^*$; in both cases, however, there is an underestimation of the shift, especially for $n \rightarrow \pi^*$ transition. We have however to observe that if we consider that rotation around $\psi(HC_mCN)$ is possible (and almost free as discussed in section 3) then the shift for the $n \rightarrow \pi^*$ transition becomes 0.16 eV. In any case, the possible discrepancies between computed and observed shift can be due to different reasons. First of all, the $n\pi^*$ band is more difficult to characterize experimentally as it is very low and it appears as a shoulder on the $\pi\pi^*$ band. Second, the experimental value

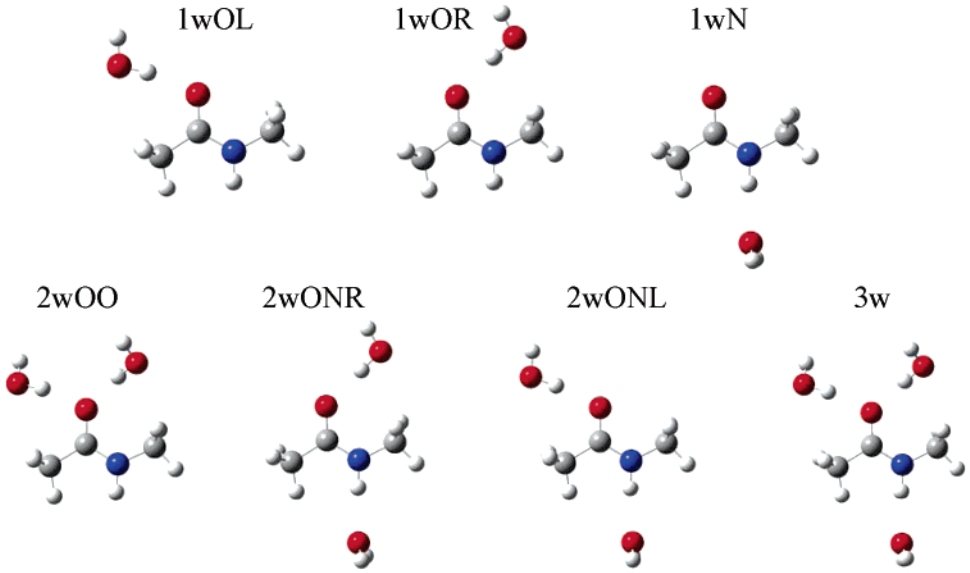

Figure 4. Schematic picture of the structures of the NMA−water clusters with indication of the corresponding labels.

TABLE 5: B3LYP/6-31+G(d,p) Optimized Geometries of NMA-1Water Clusters in the Gas Phase (vac) and in the IEF Continuum

	1wOL		1wOR		1wN	
	vac	IEF	vac	IEF	vac	IEF
CO	1.238	1.248	1.236	1.248	1.232	1.247
CN	1.358	1.346	1.356	1.346	1.361	1.348
NH	1.008	1.011	1.008	1.010	1.013	1.020
$NC_{m'}$	1.457	1.459	1.461	1.459	1.454	1.457
CC_m	1.515	1.513	1.517	1.513	1.521	1.517
$\psi(HC_mCN)$	5.3	1.8	0.1	0.4	0.2	0.6
$\varphi(CNC_{m'}H)$	172.0	179.0	178.9	179.9	180.0	179.8
(C)O···Hw	1.847	1.777	1.866	1.816		
(N)H···Ow					2.042	1.944

reported for the gas phase in fact has been measured in cyclohexane. Finally, the IEF results do not include explicit hydrogen bonding effects and these can differentially affect the two transitions.

5. H-Bonding Effects: The Supermolecule Approach

The previous analyses on IR (and UV) spectra have pointed out the need for including specific effects due to H-bonds between water and NMA. The most straightforward way to obtain this is to compute solute−solvent clusters including one, two, and three water molecules. For NMA, three waters are enough to saturate the three H-bonding donor and acceptor sites, namely two on (C)O and one on (N)H.

The clusters studied in this work are reported in Figure 4 together with the labels we shall use in the following.

5.1. Geometries. In Tables 5−6, we report a selection of geometrical parameters computed with and without the IEF continuum for the different H-bonded clusters of increasing dimension. In principle, the basis set superposition error (BSSE) can affect the calculated properties of weakly bound systems. However, Rablen et al. studied hydrogen-bonded complexes of small organic molecules and reported that, for the same basis set, BSSE in the DFT calculation is much smaller than at the MP2 level.[79] To check the validity of this assumption for the present clusters, we have repeated geometry optimizations of 2wOO, 2wON, and 3w clusters introducing the counterpoise (CP) method[80] to treat BSSE effects: the H-bond distances including CP corrections are also reported in the Tables.

The first result to remark on is that the inclusion of H-bonded water molecules makes the structure IV the most stable one both in the gas phase and within the continuum; that is, the effects of explicit H-bonds lead to the same conformer as the solvation with the continuum only (see Table 2). Other significant changes concern, once again, CO bond length which increases going from the monomer to the clusters, from smaller to larger clusters and from VAC to IEF clusters. It is interesting to note that, even if the 2wOO clusters already saturate the O H-bonding capabilities, there is a nonnegligible increase of CO bond length passing to 3w cluster: this result indicates the strong dependence of HN and CO parts.

Note also that the effects of CP corrections on clusters are negligible on the NMA geometrical parameters (and in fact they are not reported in the table) while they always lead to longer H-bond distances; in particular, the largest effect is on the (N)H···OH_2 distance which increases by 0.05 Å.

Finally, we note that H-bond distances computed for both (C)O and (N)H are in good agreement with the picture obtained from MD simulations; the analysis of RDF reported in section 1 has in fact shown a very well-defined first peak centered at 1.83 Å for (C)O−Hw and a peak centered at 2.05 Å for (N)H−Ow. These results reinforce the ability of the chosen force field to describe the solute−solvent interactions.

5.2. IR Frequencies. To analyze H-bonding effects on vibrations, we have repeated the vibrational analysis previously made for the four conformers in the gas phase and with the IEF continuum also for the NMA-nw clusters. In Table 7, we report the frequencies of the AI-AIII and NH bands computed for the different NMA-nw clusters in the gas phase and with the IEF continuum. For the AI band, two (or three) frequencies are reported, those in parentheses refer to bending motions of water molecules hydrogen bonded to amide CO and partially coupled with CO stretching.[63] For 2wOO, 2wONL, and 3w isolated clusters, we also report the frequencies obtained including counterpoise corrections (CP).

Before starting the analysis of the results, we note that CP corrections are small for all frequencies (especially for AI (CO) for which they lead to a 2 cm^{-1} change in all clusters). The main effect is on the NH stretching in the 2wON and 3w clusters: for these cases, the CP frequencies are ∼10 cm^{-1} higher than the uncorrected values. These results are not unexpected if we consider the CP effects on geometries (Table

TABLE 6: B3LYP/6-31+G(d,p) Optimized Geometries of NMA−*n*Water Clusters (*n* = 2,3) in the Gas Phase (vac) and in the IEF Continuum[a]

	2wOO		2wONL		2wONR		3w	
	vac	IEF	vac	IEF	vac	IEF	vac	IEF
CO	1.246	1.254	1.241	1.250	1.240	1.252	1.250	1.258
CN	1.349	1.341	1.352	1.340	1.352	1.342	1.344	1.336
NH	1.009	1.011	1.014	1.020	1.014	1.020	1.015	1.022
$NC_{m'}$	1.463	1.460	1.456	1.450	1.459	1.458	1.460	1.459
CC_m	1.513	1.511	1.516	1.513	1.518	1.513	1.513	1.511
$\psi(HC_mCN)$	7.7	1.2	2.2	0.2	0.1	0.2	5.1	0.4
$\varphi(CNC_{m'}H)$	175.9	175.1	180.0	179.6	179.9	179.9	177.2	174.8
(C)O···HwR	1.887	1.837			1.846	1.803	1.867	1.830
CP	1.921						1.901	
(C)O···HwL	1.872	1.805	1.827	1.769			1.850	1.806
CP	1.900		1.848				1.877	
(N)H···Ow			2.017	1.930	2.022	1.931	2.004	1.904
CP			2.073				2.051	

[a] For gas-phase 2wOO, 2wONL, and 3w also, countepoise (CP) corrected H-bond distances are reported.

TABLE 7: B3LYP/6-31+G(d,p) Frequencies of NMA-*n*w Clusters in the Gas Phase (vac) and with the IEF Continuum[a]

	AI				AII			
	vac	shift	IEF	shift	vac	shift	IEF	shift
1wOL	(1649),1721	−30	(1635),1654	−97	1563	−3	1572	+6
1wOR	(1649),1728	−23	(1631),1657	−94	1563	−3	1570	+4
1wN	1739	−12	1661	−90	1576	+10	1584	+18
2wOO	(1634,1645),1697	−54	(1624,1631),1642	109	1576	+10	1584	+18
CP	(1635,1644) 1697				1575			
2wONL	(1657),1712	−39	(1637),1651	−100	1588	+22	1609	+43
CP	(1656) 1712	(")		(")	1586	+20		(+41)
2wONR	(1656),1718	−33	(1633),1651	−100	1591	+25	1609	+43
3w	(1642,1653),1688	−63	(1623,1631),1640	−111	1603	+37	1626	+60
CP	(1641,1652) 1690	−61			1600	+34		
exp	1626 [−98]				1582 [+82/+75]			

	AIII				NH			
	vac	shift	IEF	shift	vac	shift	IEF	shift
1wOL	1297	+24	1313	+40	3663	+10	3679	+26
1wOR	1297	+24	1311	+38	3662	+9	3673	+20
1wN	1304	+31	1330	+55	3588	−65	3465	−188
2wOO	1310	+37	1321	+46	3660	+7	3641	−12
CP	1309	+36		(+45)	3658	+5		(−14)
2wONL	1324	+51	1343	+63	3575	−78	3455	−198
CP	1320	+47		(+59)	3583	−70		(−190)
2wONR	1322	+49	1340	+60	3572	−81	3454	−199
3w	1337	+64	1352	+79	3558	−95	3426	−227
CP	1333	+60		(+75)	3568	−85		(−217)
exp	1316 [+51]				3314 [−162/−146][b]			

[a] The experimental values refer to aqueous solution whereas the values in square parentheses indicates tha gas-to-water shift. The calculated shifts are with respect to the values computed for the most stable conformer of the isolated NMA (see bold values in Table 3). For the AI band, frequencies reported in parentheses refer to bending motions of water molecules hydrogen bonded to amide CO and partially coupling with CO stretching. For 2wOO, 2wONL, and 3w clusters, we also report the frequencies obtained including counterpoise corrections (CP): for IEF, the CP shifts are estimated assuming the same value of the correction calculated in the isolated clusters. [b] The shifts refer to CCl_4 and CH_2Cl_2 solutions, respectively.[66]

6): the main changes are in fact on the (N)H···Ow H-bond distances which become longer when CP corrections are considered.

In the previous comparison between isolated and IEF-solvated NMA (see Table 2), we have analyzed the "bulk" electrostatic effects (namely those described by the IEF continuum model alone) on the four amide modes, and here we can extend the analysis by introducing H-bonding effects (through gas-phase clusters) and the combination of the two effects (through IEF clusters).

Even if of very different nature in the different modes, in all cases, H-bonding effects alone poorly describe the observed gas-to-water shifts and also when all H-bonding sites are saturated as in the largest 3w cluster, the shifts calculated with respect to the isolated NMA are still too small (with the only exception of AIII mode). Only by adding the effects of IEF continuum, the computed shifts can be successfully compared to the experimental ones showing that both short-range (or first-shell) and bulk (or second and higher shells) effects are necessary for a correct description of the vibrational spectra of NMA in water.

This is not an unexpected result as previous computational studies already showed that, at least for the amide I band, water molecules in the second hydration shell are important.[69,70] Here, however, we can further extend this analysis and check if these bulk effects act as an additive term or if, conversely, they synergically combine with the shorter and more specific H-bonding effects.

From the results of Table 7, it is evident that IEF continuum effects (i.e., shifts) are different for the different clusters (also within the same nw set) and for the different modes.

TABLE 8: Mean Residence Times (in ps)

group	$t^* = 0$ ps	$t^* = 2$ ps
C═O	1.9 ± 0.5	3.2 ± 0.1
N−H	0.7 ± 0.5	2.2 ± 0.1

For the amide I (CO) band, the addition of the bulk effect through the IEF continuum leads to a significant improvement in the agreement with the experiments with respect to the isolated clusters, namely we find 3-times larger shifts for the various clusters with the exception of the largest one in which the shift is less than 2 times larger. In any case, it appears evident that both first and more outer solvation shells have to be taken into account to correctly describe the observed shift in the CO band. Going into more details, the best agreement between calculated and experimental shifts is found for the 2wON(L or R) solvated clusters for which we compute a shift of 100 cm^{-1}, e.g., almost exactly the experimental observation. It is also important to note that two H-bonds on the carbonyl oxygens lead to an overestimation of the shift both in the 2w and in 3w clusters.

In the case of the NH band, the continuum alone gives a wrong description (see Table 2). It comes out evident that this wrong description was due to a lack of H-bond effects; in fact as soon as we introduce an H-bonded water to (N)H, the frequency of the band is significantly reduced also in the isolated clusters. As for the amide I, also here we need to introduce an external continuum to reach a sufficient agreement with the experiments. In this case, however, the nature (or the geometry) of the cluster becomes important and the differences among the different clusters are amplified by the continuum. This shows that when H-bonding effects are dominant their combination with long-range interactions (i.e., the continuum) becomes important too, making the two effects synergic. The 2wON clusters which gave the best results for amide I band here seem to overestimate the effect. However, it has to be noted that the experimental shift does not really refer to gas-phase measurements but to apolar solvents: we cannot thus use it as a real reference but instead as a an upper limit, and thus the computed 198 cm^{-1} shift can be considered a good description.

For the two AII−AIII complementary bands, the analysis is more difficult. Obviously, both modes require us to introduce the (N)H···OH_2 interaction (and thus 1wN, 2wON, and 3w clusters are the best candidates). However, the results show that these modes, due to their hybrid nature, are affected by all of the other interactions in a complex way. In addition, for the two modes, there is the further compliance that they are strongly affected by the methyl rotations, and thus, the definition of the reference gas-phase value is not straightforward. What we have done is to consider, as reference, the value of the frequency corresponding to the most stable conformer in the gas phase; this surely works well for CO band for which the effect of rotations are negligible but also for NH for which the value does not change if we consider an average of the different conformers. For AII and AII however, the analysis in terms of different conformers would reduce the reference values by ∼10 cm^{-1} for AII and increase it by ∼5 cm^{-1} for AIII: this means that all of the computed shifts reported in Table 8 should be increased by 10 cm^{-1} for AII and decreased ca. 5 cm^{-1} for AIII. In this way, the best description in terms of the IEF 2wON clusters is obtained, reaching a perfect agreement with experiments for AIII and still good but underestimated shift for AII. For this latter, we have also to note that the experimental data are very "dispersed".

As a final note, we observe that the nwO clusters present the expected splitting of the AI band due to the mixing with HOH bending vibration of the hydrogen-bonded water molecules. The computed splitting represents a further element in favor of the mixed continuum/H-bond description; in gas-phase clusters in fact, we find a too large splitting (70 cm^{-1} for 1wO clusters which reduces to 50 cm^{-1} in 2w and 40 cm^{-1} in 3w clusters) while adding the IEF continuum this reduces to 20 cm^{-1} in all different clusters exactly as experimentally observed.[63]

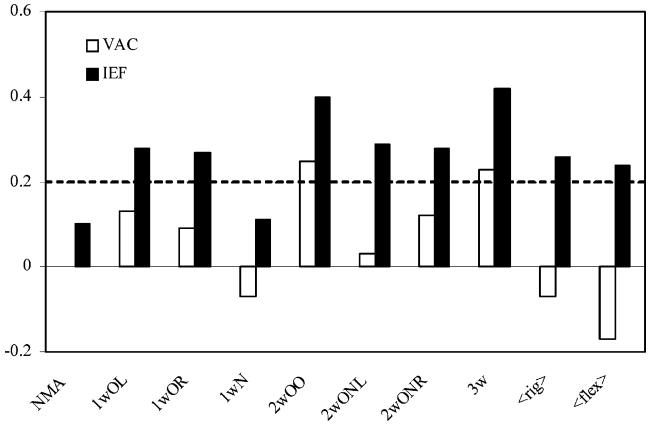

Figure 5. Analysis of absorption energies for the $n \rightarrow \pi^*$ transitions in the NMA−water QM and MD clusters in terms of the corresponding shifts with respect to the isolated NMA: as reference excitation energies of isolated NMA, we have considered those obtained for the most stable conformer (II). Notice that the shifts obtained with a continuum-only model is also reported. Dotted line shows the experimental value (see text for details).

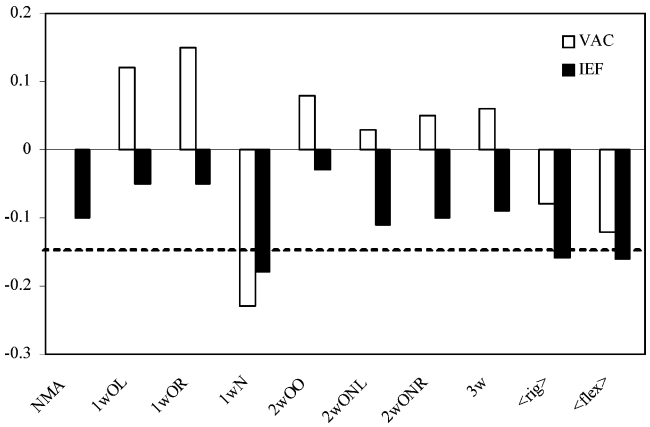

Figure 6. Analysis of absorption energies for the $\pi \rightarrow \pi^*$ transitions in the NMA−water QM and MD clusters in terms of the corresponding shifts with respect to the isolated NMA: as reference excitation energies of isolated NMA, we have considered those obtained for the most stable conformer (II). Notice that the shifts obtained with a continuum-only model is also reported. Dotted line shows the experimental value (see text for details).

5.3. Absorption Energies. The same analysis in terms of bulk and H-bonding effects made for IR frequencies is repeated for electronic transitions.

To have a more direct picture of solvent effects in Figures 5 and 6, we present the analysis of absorption energies for the $n \rightarrow \pi^*$ and $\pi \rightarrow \pi^*$ transitions in terms of the corresponding shifts with respect to the isolated NMA: as reference excitation energies of isolated NMA, we have considered those obtained for the most stable conformer (II), whereas the dotted lines indicate the experimental data (namely +0.2 eV for $n \rightarrow \pi^*$ and −0.14 eV for $\pi \rightarrow \pi^*$). We note that in these graphs we have also reported the shifts obtained with a continuum-only model that we already discussed in section 4.2.

In addition to the NMA-nw clusters used in the previous section (see also Figure 4), here we also consider an alternative

set of clusters obtained from MD simulations as described in section 2.2. The comparison of the two alternative sets of clusters (either QM or MD) to represent the effects of the first solvation shell(s) should give further hints on the nature of the solvation around NMA, and in particular, it should help in analyzing if these effects are better represented in terms of rigid structures obtained as minima of the potential energy surfaces, or if, by contrast, dynamics, and therefore a variety of different and representative structures, are needed to properly describe the hydration effects.

For MD-derived clusters, two sets of data are reported, each one referring to the type of model, rigid or flexible, used for NMA in the MD simulations. We recall that in the rigid model the NMA structure is kept frozen in the IV-like (or solution) structure obtained from the previous QM-IEF optimization, whereas in the flexible model, all internal degrees of freedom can change. For both models, the value reported in Figures 5 and 6 is obtained as an average on a set of 60 (for the flexible) and 40 (for the rigid) different clusters (see section 2.2 for more details) with the addition of an external continuum solvent mimicking the bulk water.

The analysis of results of Figures 5−6 involves different aspects. First, let us only consider the effects of H-bonds by analyzing isolated NMA-nw clusters (VAC in the figures). H-bonding effects on the two transitions are opposite if the H-bond acts on (C)O or on (N)H. In fact, including one or two water molecules close to carbonyl oxygen induces a blue shift on both transitions, whereas when only the (N)H−Ow bond is included (1wN), the total effect is a red-shift instead of a blue shift; such a red shift is quite small in the case of the $n \rightarrow \pi^*$ transition, whereas it is significant for the $\pi \rightarrow \pi^*$ one. When the two different H-bonds are combined (2wON and 3w clusters), the global effect is again a blue-shift but smaller than in the 1(2)wO(O) cases. In the case of the $\pi \rightarrow \pi^*$ transition also, the largest 3w cluster presents a very small shift due to the two opposite effects of the two H-bonds that almost cancel out. These findings can be used to analyze the different models adopted.

The inclusion of IEF improves the description leading to the correct sign in the shift for all of the possible clusters. For both transitions, the long-range electrostatic effects of the bulk solvent appear to be more important than the local specific H-bond as expected from the previous good behavior of the continuum only model (see Table 4). The set of IEF results however, gives additional information. First, $n \rightarrow \pi^*$ results show that summing the effects of two H-bonds on carbonyl oxygen together with the long-range interactions leads to an overestimation of the observed solvent-effect (around twice the experimental 0.2 eV value); much better descriptions are given by the two 1wO clusters and by the mixed 2wON clusters (even if slightly overestimated in this latter case). The 1wN result remains too low, and in practice, it does not add anything more to the continuum-only result.

Concerning the $\pi \rightarrow \pi^*$ transition, the best results are represented by the 2wON clusters, but in any case, we do not observe a significant improvement with respect to the continuum-only model. Once again by assuming two (C)O···HOH interactions (2wOO), we obtain a noncorrect description even if the inclusion of the continuum partially improves the completely wrong picture obtained in the isolated clusters. Finally, the 1wN cluster leads to a slightly too large effect.

If we pass from the static picture of QM-optimized clusters to that of statistical type obtained in terms of MD clusters and we also include bulk effects with the external continuum, a very good result is obtained for $\pi \rightarrow \pi^*$ transition, whereas for $n \rightarrow \pi^*$, we find a slightly too large shift. This result however has to be reappraised if we consider that the experimental shift refers to cyclohexane and not to the gas phase and thus it has to be read as a lower limit. We point out that the same clusters without the continuum give a negligible shift for $n \rightarrow \pi^*$ and a too small one for $\pi \rightarrow \pi^*$. It is also interesting to note that when the same clusters are described at a simpler level, namely by representing the water molecules in terms of TIP3P[42] classical charges, the description we obtain for the two transitions is different: although the $n \rightarrow \pi^*$ transition is reasonably described (0.1 eV shift), the $\pi \rightarrow \pi^*$ one is not reproduced at all (small and wrong sign shift). Therefore, only pure electrostatic effects are not enough to describe the processes and a more complex picture of the interactions involving a quantum mechanics electronic solvent description is needed.

These results strongly indicate that, as far as UV absorption spectra concerns, short-range and long-range solvation effects are required in order to properly describe the observed shifts. For short-range effects, a better picture is obtained by including explicit solvent molecules (namely those representing the first solvation shells) and mediating on different configurations (i.e., introducing a statistical description) while long range effects are properly and effectively described by adding a reaction field through a continuum model.

6. Further MD Analysis

With the aim of a better understanding of the results presented in the previous sections, a deeper analysis of the MD simulations have been performed. In this sense, the hydration structure adopted by the solvent around the NMA molecule have been studied by means of spatial distribution functions (SDF)[81] and coordination numbers (CN). The SDF 3-D pictures, defined by isovalue surfaces, show the highest probability regions for finding a given type of atom in the neighborhood of a site of interest. In Figure 7, the higher water oxygen and hydrogen density regions around the amide group can be observed.

It is interesting to notice how the structure adopted by the water molecules is clearly better defined in the case of the carbonyl group, in agreement with the RDFs previously discussed: water oxygen and hydrogen atoms show well separated sphere-like distributions not sharing common space regions and running almost in parallel (concentric shells). Interestingly the lowest probability of finding a water molecule atoms is found along the C═O bond direction. This fact is easily understood if, on average, two water molecules hydrate the carbonyl group simultaneously: the average Ow−O−Ow angle is 93° while 174° is obtained for the dihedral Ow1−O═C−Ow2, that is, the two water oxygens and the C═O unit prefer a quasi planar arrangement, minimizing in this way the water−water repulsion when both solvent molecules hydrate the carbonyl group. However, the angle probability distribution (not shown) reflects an interesting fact that can be clearly observed in the SDF: despite this preference for planar arrangements, water molecules occupy a wide range of space around the carbonyl group, allowing for many different relative positions with respect to the NMA peptide plane.

In the N−H case, the image that SDF supports is different. In this case, the lobe distributions adopted by both types of water atoms show how the water orientation is not as rigidly defined as in the CO case. Nonetheless, the motion of the water molecule hydrating the amide group is much more hindered than in the carbonyl case, sampling preferentially the peptide plane. This fact is only reflecting the directionality present in the C═O···OH_2 hydrogen bond interaction.

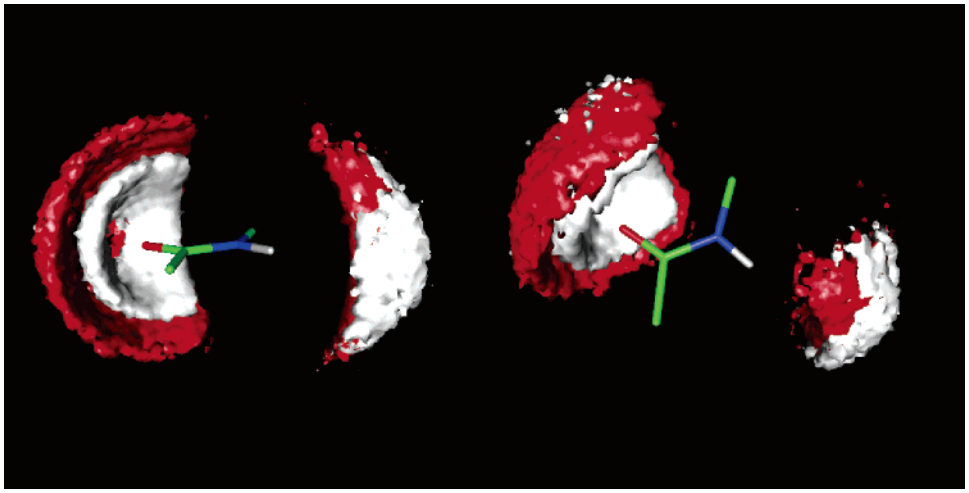

Figure 7. Water oxygen (red) and hydrogen (white) spatial distribution functions around amide and carbonyl groups contoured at 75 above bulk density.

The integration numbers previously obtained from the RDFs (see section 1) reflect a dynamic equilibrium between different coordination numbers. Using the geometrical criterion based on the O−OW and H−OW radial distribution functions, the solvation coordination number distributions for the C═O and N−H groups cover, with nonnegligible contributions, the ranges 1−3 and 0−2, respectively. This feature can be correlated with dynamic information of the water molecules hydrogen bonded to the amide group, in particular the so-called mean residence times (τ) that have been computed using the method of Impey et al.[82] (see Table 8). Within this model, the parameter t^* defines the maximum transient period a solvent molecule can leave the considered region, in this case the hydration shell defined by the corresponding (C)O−OW and (N)H−OW RDFs, without losing its ascription to it. Values of 0 ps[83] and 2 ps[82] have been frequently used in the literature for aqueous solutions and are applied in this work as well.

As can be observed, no large differences are found in the residence times of water molecules interacting with the two groups defining the peptide unit. Anyhow, the solvent molecules in the vicinity of the carbonyl group show a clear tendency to present higher τ values reflecting a stronger solute−solvent interaction. At first sight, these results could seem unexpected if the much higher definition of the carbonyl hydration shell, according to the RDFs and SDFs, is taken into account. However, the parameter that determines to a large extent the τ values is the value of the RDF at the minimum between the first and the second shells. According to that, the differences between the (C)O−OW and (N)H−OW RDFs are relatively small. Finally, it is worth pointing out that these results are similar to the residence times found for pure water simulations.[82]

The image obtained from the MD structural and dynamic analysis of the peptide group hydration, although strongly supports the idea of one molecule hydrating the amide group and two solvent molecules, in the case of the carbonyl one, is far from the static quantum mechanics view obtained from the structure outcome of the energy minimization process. Although the solute−solvent kind of interactions are well defined at QM and MM levels, the solvent distribution, as shown by the SDFs, is much more spread out in space. In the case of the C═O unit, this fact can, at least partially, explain the overestimation observed in the IR shift for the amide I normal mode due to solvent effects when, on top of the continuum, two water molecules H-bonded to the carbonyl group are considered. The QM geometry based approach overestimates the effect because only the most energetically favorable arrangement is considered for the property calculation. However, other still two-water-like structures with rather different orientations sampled at room temperature are omitted.

7. Conclusions

This paper, together with the parallel one on the NMR shifts in water and acetone,[33] addresses two basic aspects of the study of peptides in solution: the pragmatic aspect of how to describe, with a reliable but computationally feasible procedure, the solvent effects on different spectroscopic properties, and the general problem of how to asses the limitations and potentialities of different theoretical models in the description of the physics and the dynamics of the solvation of peptides in protic and nonprotic polar solvents.

Clearly, the two problems are not only strictly linked but they have to be addressed in parallel if a correct picture is pursued. In this sense, we have shown that this kind of analysis can be obtained by using various computational tools and combining them with experimental measurements of different nature. Only the consideration of spectroscopic properties involving different physical processes, and thus differently affected by the environment, can in fact allow a more complete description of the solvation phenomenon. In a disordered and dynamic medium, as a liquid, in fact, the interactions among the components can be of very different character, either acting more locally or, by contrast, extending farther.

For nonlocal interactions, it is nowadays well established that an effective strategy is represented by continuum solvation models if these latter properly include accurate descriptions of the solute−solvent interactions by using molecularly shaped cavities and possibly account for solvent dynamic effects (like here nonequilibrium effects).

Concerning local interactions, the picture becomes much more complex. A useful simplification is that of introducing another distinction between labile and persistent interactions, based on the intrinsic time necessary to define and to measure a property:[84] labile interactions must be averaged, persistent interactions can be taken as they are. The present study of hydration of NMA, however, seems to show that this phenomenon is much more complex than usually thought and that this simplification can lead to erroneous conclusions.

Generally, it is in fact assumed that hydrogen bonds between water and the peptide group could be defined as a local

interaction, persistent over time scales covering many types of measurements. This picture is also confirmed by the analysis of RDF functions one obtains from MD or Monte Carlo simulations: as reported in section 2.2 in fact, the RDFS show, for O−HW, a very well-defined first peak that integrates to about two water molecules and, for H−OW, a less defined peak integrating to one. The average image emerging from these figures is that two hydrogen atoms of two different water molecules hydrate the carbonyl group and another water molecule hydrates the amide hydrogen. On the basis of this picture, one might guess that a good prediction of the spectroscopic properties of NMA in water is obtained in terms of clusters of NMA and three water molecules, two H-bonded to the carbonyl oxygen and one to the amide hydrogen. By contrast, from the analysis of IR and UV spectra reported in the previous sections, it follows that this picture is not completely right, or at least it is incomplete.

The IR analysis shows that the supermolecule approach is needed to predict the blue or red character of the solvent induced frequency shifts. However, a better agreement is found when the continuum model is added on top of the aggregates containing the solute and some explicit solvent molecules. In this case, a static picture containing the minimum number of solvent molecules to saturate H-bonds seems to grasp most of the nature of the vibrational process. Nonetheless, the observed tendency shows an overestimation in three of the four normal modes analyzed. Likely, the QM geometry based approach overestimates the solvent effect because only the most energetically favorable geometry is considered in the property calculation. However, other structures, still involving two waters but with rather different orientations sampled at room temperature (as inferred from the SDF analysis) are omitted. In this sense, corrections to the frequency shifts as a function of the solvent molecule positions[71] should consider not only the distance to the carbonyl and amide groups but also the angular dependence.

Concerning the UV/VIS electronic transitions, the need of including long-range effects together with explicit water molecules in the quantum calculations is clear. In this case, the static picture taken from the QM optimized structures fail to reproduce the property, even at a qualitative level. Therefore, a more "disordered" view, accounting for the statistical nature of the solute−solvent interactions is compulsory. The effect of the continuum must be considered if a quantitative agreement with the experimental values is pursued. In this sense, this approach supposes an effective way of including the electrostatic long-range effects. An alternative approach found in the literature[49] producing similar results considers explicitly solvent molecules belonging to the second and outer solvation shells in the QM calculations. It is easy to understand that larger number of solvent molecules implies, because of the disordered nature of the solvent, a larger the number of cluster calculations if a proper sampling is performed (convergence), and, at the same time, more expensive QM computations limiting the level of theory employed for them. The PCM methodology clearly avoids these limitations.

Acknowledgment. J.M.M. thanks Junta de Andalucía and Spanish DGICYT (BQU2002-02217) for financial support.

References and Notes

(1) Burling, F. T.; Weis, W.; Flaherty, K.; Brünger, A. *Science* **1996**, *271*, 72.

(2) Grant, E.; McClean, V.; Nightingale, N.; Sheppard, R.; Chapman, M. *Bioelectromagnetics* **1986**, *7*, 151.

(3) Otting, G.; Liepinsh, E.; Wuthrich, K. *Science* **1991**, *254*, 974.

(4) Pal, S.; Peon, J.; Zewail, A. *Proc. Natl. Acad. Sci.* **2002**, *99*, 1763.

(5) Woutersen, S.; Mu, Y.; Stock, G.; Hamm, P. *Proc. Natl. Acad. Sci.* **2001**, *98*, 11254.

(6) Gregory, R. *Protein−Solvent Interactions*; Dekker: New York, 1995.

(7) Orozco, M.; Luque, F. *Chem. Rev.* **2000**, *100*, 4187.

(8) Guo, H.; Karplus, M. *J. Phys. Chem.* **1992**, *96*, 7273.

(9) Serrano-Andres, L.; Fulscher, M. *J. Am. Chem. Soc.* **1996**, *118*, 12190.

(10) Han, W.-G.; Suhai, S. *J. Phys. Chem.* **1996**, *100*, 3942.

(11) Markham, L. M.; Hudson, B. S. *J. Phys. Chem.* **1996**, *100*, 2731.

(12) Szalay, P.; Fogarasi, G. *Chem. Phys. Lett*. **1997**, *270*, 406.

(13) (a) Hirst, J. D.; Hirst, D. M.; Brooks, C. L., III. *J. Phys. Chem. A* **1997**, *101*, 4821. (b) Watson, T. M.; Hirst, J. D. *J. Phys. Chem. A* **2002**, *106*, 7858. (c) Watson, T. M.; Hirst, J. D. *J. Phys. Chem. A* **2003**, *107*, 6843.

(14) Samdal, S. *J. Mol. Struct*. **1998**, *440*, 165.

(15) Villani, V.; Alagona, G.; Ghio, C. *Mol. Eng*. **1999**, *8*, 135.

(16) Kang, Y. K. *J. Phys. Chem. B* **2000**, *104*, 8321.

(17) Kubelka, J.; Keiderling, T. *J. Phys. Chem. A* **2001**, *105*, 10922.

(18) Cuevas, G.; Renugopalakrishnan, V.; Madrida, G.; Haglerd, A. T. *Phys. Chem. Chem. Phys.* **2002**, *4*, 1490.

(19) Nandini, G.; Sathyanarayana, D. N. *J. Mol. Struct. (THEOCHEM)* **2002**, *579*, 1.

(20) Mannfors, B. E.; Mirkin, N. G.; Palmo, K.; Krimm, S. *J. Phys. Chem. A* **2003**, *107*, 1825.

(21) Papamokos, G. V.; Demetropoulos, I. N. *J. Phys. Chem. A* **2004**, *108*, 7291.

(22) Zhang, R.; Li, H.; Lei, Y.; Han, S. *J. Mol. Struct*. **2004**, *693*, 17.

(23) Bour, P. *J. Chem. Phys.* **2004**, *121*, 7545.

(24) Tomasi, J.; Persico, M. *Chem. Rev.* **1994**, *94*, 2027.

(25) Cramer, D. Truhlar, *Chem. Rev.* **1999**, *99*, 2161.

(26) Tomasi, J.; Cammi, R.; Mennucci, B.; Cappelli, C.; Corni, S. *Phys. Chem. Chem. Phys.* **2002**, *4*, 5697.

(27) Li, J.; Cramer, C. J.; Truhlar, D. G. *Int. J. Quantum Chem*. **2000**, *77*, 264.

(28) Mennucci, B.; Martínez, J. M.; Tomasi, J. *J. Phys. Chem. A* **2001**, *105*, 7287.

(29) Mennucci, B. *J. Am. Chem. Soc.* **2002**, *124*, 1506.

(30) (a) Cossi, M.; Crescenzi, O. *Theor. Chem. Acc*. **2004**, *111*, 162. (b) Cossi, M.; Crescenzi, O. *J. Chem. Phys.* **2003**, *118*, 8863. (c) Aquilante, F.; Cossi, M.; Crescenzi, O.; Scalmani, G.; Barone, V. *Mol. Phys.* **2003**, *101*, 1945.

(31) (a) Cancès, E.; Mennucci, B. *J. Math. Chem*. **1998**, *23*, 309. (b) Cancès, E.; Mennucci, B.; Tomasi, J. *J. Chem. Phys.* **1997**, *107*, 3031. (c) Mennucci, B.; Cancès, E.; Tomasi, J. *J. Phys. Chem. B* **1997**, *101*, 10506.

(32) (a) Miertus, S.; Scrocco, E.; Tomasi, *J. Chem. Phys.* **1981**, *55*, 117. (b) Cammi, R.; Tomasi, J. *J. Comput. Chem*. **1995**, *16*, 1449.

(33) Mennucci, B.; Martínez, J. M. *J. Phys. Chem. B* **2005**, *109*, 9830.

(34) (a) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648. (b) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.

(35) Bondi, A. *J. Phys. Chem.* **1964**, *68*, 441.

(36) Mennucci, B.; Cammi, R.; Tomasi, J. *J. Chem. Phys.* **1998**, *109*, 2798.

(37) Cammi, R.; Mennucci, B.; Tomasi, J. *J. Phys. Chem. A* **2000**, *104*, 5631.

(38) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*, revision A.1; Gaussian, Inc.: Pittsburgh, PA, 2003.

(39) Smith, W.; Leslie, M.; Forester, T. R. *CCLRC*; Daresbury Laboratory: Daresbury, Warrington, U.K..

(40) Cornell, W. D.; Cieplak, P.; Bayly, C. Y.; Gould, I. R.; Merz, K. M., Jr.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.

(41) Kaminski, G.; Jorgensen, W. L. *J. Phys. Chem.* **1996**, *100*, 18010.

(42) Jorgensen, W. L.; Chandreskhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L *J. Chem. Phys.* **1983**, *79*, 926.

(43) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577.

(44) Zhang, R.; Li, H.; Lei, Y.; Han, S. *J. Phys. Chem. B* **2004**, *108*, 12596.
(45) Rick, S. W.; Berne, B. J. *J. Am. Chem. Soc.* **1996**, *118*, 672.
(46) Gao, J.; Freindorf, M. *J. Phys. Chem. A* **1997**, *101*, 3182.
(47) Beglov, D.; Roux, B. *J. Phys. Chem. B* **1997**, *101*, 7821.
(48) Buck, M.; Karplus, M. *J. Phys. Chem. B* **2001**, *105*, 11000.
(49) Rocha, W.; Almeida, K. D.; Coutinho, K.; Canuto, S. *Chem. Phys. Lett.* **2001**, *345*, 171.
(50) Iuchi, S.; Morita, A.; Kato, S. *J. Phys. Chem. B* **2002**, *106*, 3466.
(51) Ataka, S.; Takeuchi, H.; Tasumi, M. *J. Mol. Struct.* **1984**, *113*, 147.
(52) Spencer, J.; Berger, S.; Powell, C.; Henning, B.; Furman, G.; Loffredo, W.; Rydgerg, E.; Neubert, R.; Shoop, C.; Blauch, D. *J. Phys. Chem.* **1981**, *85*, 1236.
(53) Radzicka, A.; Wolfenden, R. *Biochemistry* **1988**, *27*, 1664.
(54) Mirkin, N. G.; Krimm, S. *J. Am. Chem. Soc.* **1991**, *113*, 9742.
(55) Ohashi, N.; Hougen, J. T.; Suenram, R. D.; Lovas, F. J.; Kawashima, Y.; Fujitake, M.; Pyka, J. *J. Mol. Spectrosc.* **2004**, *227*, 28.
(56) Katz, J.; Post, B. *Acta Crystallogr.* **1960**, *13*, 624.
(57) Hagler, A.; Leiserowitz, L.; Tuval, M. *J. Am. Chem. Soc.* **1976**, *98*, 4600.
(58) Balazs, A. *J. Mol. Struct.* **1987**, *153*, 103.
(59) Mirkin, N. G.; Krimm, S. *J. Mol. Struct. (THEOCHEM)* **1995**, *334*, 1.
(60) (a) Song, S.; Asher, S.; Krimm, S.; Bandekar, J. *J. Am. Chem. Soc.* **1988**, *110*, 8547. (b) Song, S.; Asher, S.; Krimm, S.; Shaw, K. D. *J. Am. Chem. Soc.* **1991**, *113*, 1155.
(61) Mayne, L.; Hudson, B. *J. Phys. Chem.* **1991**, *95*, 2962.
(62) Wang, Y.; Purrello, R.; Ceorgiout, S.; Spiro, T. G. *J. Am. Chem. Soc.* **1991**, *113*, 6359; 6368.
(63) Chen, X. G.; Schweitzer-Stenner, R.; Krimm, S.; Mirkin, N. G.; Asher, S. *J. Am. Chem. Soc.* **1994**, *116*, 11141.
(64) Chen, X. G.; Schweitzer-Stenner, R.; Asher, S. A.; Mirkin, N. G.; Krimm, S. *J. Phys. Chem.* **1995**, *99*, 3074.
(65) Mirkin, N. G.; Krimm, S. *J. Mol. Struct.* **1996**, *377*, 219.
(66) Herrebout, W.; Clou, K.; Desseyn, H. *J. Phys. Chem. A* **2001**, *105*, 4865.
(67) Schweitzer-Stenner, R. J. *Raman Spectosc.* **2001**, *32*, 711.
(68) Kwac, K.; Cho, M. *J. Chem. Phys.* **2003**, *119*, 2247; Ibiden 2256.
(69) Bour, P.; Keiderling, T. A. *J. Chem. Phys.* **2003**, *119*, 11253.
(70) Torii, H. *J. Phys. Chem. A* **2004**, *108*, 7272.
(71) Besley, N. A. *J. Phys. Chem. A* **2004**, *108*, 10794.
(72) See, for example: (a) Surewicz, W. K.; Mantsch, H. H. *Biochim. Biophys. Acta* **1988**, *952*, 115. (b) Braiman, M. S.; Rothschild, K. J. *Annu. Rev. Biophys. Chem.* **1998**, *17*, 541.
(73) (a) Pople, J. A.; Scott, A. P.; Wong, M. W.; Radom, L. *Israel J. Chem.* **1993**, *33*, 345. (b) Rauhut, G.; Pulay, R. *J. Phys. Chem.* **1995**, *99*, 3093. (c) Scott, A. P.; Radom, L. *J. Phys. Chem.* **1996**, *100*, 16502. (d) Fast, P. L.; Corchado, J.; Sánchez, M. L.; Truhlar, D. G. *J. Phys. Chem. A* **1999**, *103*, 3139.
(74) Cappelli, C.; Silva, C. O.; Tomasi, J. *J. Mol. Struct. (THEOCHEM)* **2001**, *544*, 191.
(75) Kaya, K.; Nakagura, S. *Theor. Chim. Acta* **1967**, *7*, 117.
(76) Nielsen, E.; Schellman, J. *J. Chem. Phys.* **1967**, *71*, 2297.
(77) Akiyama, M.; Torii, H. *Spectrochim. Acta Part A* **2000**, *56*, 137.
(78) Köddermann, T.; Ludwig, R. *Phys. Chem. Chem. Phys.* **2004**, *6*, 1867.
(79) Rablen, P. R.; Lockman, J. W.; Jorgensen, W. L. *J. Phys. Chem. A* **1998**, *102*, 3782
(80) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1990**, *19*, 553.
(81) Kusalik, P. G.; Laaksonen, A.; Svishchev, I. M. *Molecular Dynamics: From Classical to Quantum Methods*; Elsevier: Amsterdam, 1999; Chapter 3.
(82) Impey, R. W.; Madden, P. A.; McDonald, I. R. *J. Phys. Chem.* **1983**, *87*, 5071.
(83) Garcia, A. E.; Stiller, L. *J. Comput. Chem.* **1993**, *14*, 1396.
(84) Cappelli, C.; Mennucci, B.; da Silva, C. O.; Tomasi, J. *J. Chem. Phys.* **2000**, *112*, 5382.
(85) Kitano, M.; Fukuyama, T.; Kuchitsu, K. *Bull. Chem. Soc. Jpn.* **1973**, *46*, 384.