JCTC Journal of Chemical Theory and Computation

Infrared Spectroscopy of *N*-Methylacetamide Revisited by ab Initio Molecular Dynamics Simulations

M. P. Gaigeot,*,† R. Vuilleumier,‡ M. Sprik,§ and D. Borgis†

Laboratoire de Modélisation des Systèmes Moléculaires Complexes, Université d'Evry val d'Essonne, Rue Père A. Jarland, 91025 Evry, France, Laboratoire de Physique Théorique des Liquides, Université Pierre et Marie Curie, Case Courrier 121, 4 place Jussieu, 75005 Paris, France, and Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, United Kingdom

Received February 14, 2005

Abstract: The density functional theory based molecular dynamics simulation method ("Car−Parrinello") was applied in a numerical study of the electronic properties, hydrogen bonding, and infrared spectroscopy of the trans and cis isomer of *N*-methylacetamide in aqueous solution. A detailed analysis of the electronic structure of the solvated molecules, in terms of localized Wannier functions and Born atomic charges, is presented. Two schemes for the computation of the solute infrared absorption spectrum are investigated: In the first method the spectrum is determined by Fourier transforming the time correlation function of the solute dipole as determined from the Wannier function analysis. The second method uses instead the molecular current−current correlation function computed from the Born charges and atomic velocities. The resulting spectral properties of *trans*- and *cis*-NMA are carefully compared to each other and to experimental results. We find that the two solvated isomers can be clearly distinguished by their infrared spectral profile in the 1000−2000 cm^{-1} range.

1. Introduction

Vibrational spectroscopy (infrared and Raman) can often provide information about molecular structures in the condensed phase which is complementary to or even hidden from diffraction methods. This is why the technique is commonly applied for the study of secondary structures of proteins[1−4] in solution or membranes. Structural motifs such as α-helices or β-sheets can be identified by characteristic changes in position and shape of linear infrared or Raman bands,[1−4] in particular the amide bands (mostly amide I, amide II, and amide III) are used as structural probes. Amide I gives the most sensitive response to changes in the secondary structure of proteins.[2] This is partly due to the fact that the amide I vibration appears in a spectral region with very little overlap with other internal modes of proteins and to the fact that its spectral intensity generally dominates the infrared spectrum. In addition to reflecting the molecular conformation,[2] the amide I band shape is also strongly affected by formation of intermolecular hydrogen bonds with the solvent.[2,5−7] Moreover, the frequency gap between the amide I and amide II bands represents some measure of the interaction between carbonyl and amide groups on the protein backbone. Consequently, the amide I mode is also used as a probe in nonlinear-IR-2D spectroscopy of small peptides.[8−11]

Assignment of vibrational modes is an important step in the structural and dynamical interpretation of the spectral fingerprints of conformation or solvation. Here is where modeling and computation can be most helpful, in particular when a full spectrum (including intensities) can be computed and directly compared to experimental data. As we have demonstrated in a recent study of the IR absorption of aqueous uracil[12] it is now possible to perform such a detailed calculation for small models of bio-organic molecules in finite temperature bulk solution with all water molecules treated at the same level of theory as the solute. This

* Corresponding author e-mail: gaigeot@ccr.jussieu.fr.
† Université d'Evry val d'Essonne.
‡ Université Pierre et Marie Curie.
§ University of Cambridge.

10.1021/ct050029z CCC: $30.25
Published on Web 07/07/2005

development is the result of the latest advances in density functional theory based molecular dynamics simulation ("Car−Parrinello"[13]), brought about by the implementation of the modern theory of polarization and the maximally localized Wannier functions. This opened up the field of first principle simulation to the study of infrared[14−16] and Raman[17] spectroscopy.

Following on our calculation of the infrared spectrum of aqueous uracil[12] (one of the RNA nucleic acid bases), we investigate in the present paper the infrared spectroscopy of another model bio-organic molecule, NMA, using similar ab initio molecular dynamics methods. Our main objective is again to assess whether Car−Parrinello simulations can be relied upon to give a proper account of the experimental spectrum of this elementary protein building block in its aqueous environment. Comparing the calculated infrared spectrum of gas-phase and aqueous-phase NMA, band shape modifications as well as band shifts will be correlated with the effect of the hydrogen bonding on internal motions of the molecule. Of special interest is the question of the spectroscopic characterization of the two distinct conformers. The change in structure between *trans*- and *cis*-NMA as well as an expected different organization of surrounding solvent molecules gives rise to different vibrational signatures recorded in infrared spectra. We have thus analyzed the infrared response to a change in conformation and compared the hydration properties of *trans*- and *cis*-NMA.

The paper also serves a more technical purpose, namely to present and compare two alternative ways of calculating infrared spectra of molecular solutes in a molecular solvent. (i) The first is the method applied in ref 12 which consists of calculating the individual molecular dipole moments (for the solute and each solvent molecule) by using localized Wannier functions.[12,14−16] (ii) The second approach, introduced here, is based on the calculation of dipole derivatives (or currents) through linear response theory. The total current is decomposed into solute and solvent contributions. Compared to the first scheme, based on partitioning of charge density, the second approach using the current representation is more fundamental leading to a rigorous separation of solute and solvent signal. However, requiring a coupled self-consistent perturbation calculation, the computational costs are also considerably higher.

Organization of the paper is as follows. Section 2 contains some relevant background information on NMA. Section 3 is a summary of the conditions of the Car−Parrinello molecular dynamics and the methods used in our work for the calculation of infrared spectra of solutes. In section 4 the structural properties of isolated and aqueous *trans*- and *cis*-NMA are described including details of the hydrogen bonding network between solute and solvent. The electronic properties of the solute are discussed too, in terms of localized Wannier functions, Born charges, and dipole moments. In section 6, we report our findings for the vibrational density of states (VDOS) of the various systems studied. Section 7 is devoted to the calculated infrared spectra. The comparison between the gas-phase and aqueous-phase spectral properties is discussed in detail. The last section concludes our study.

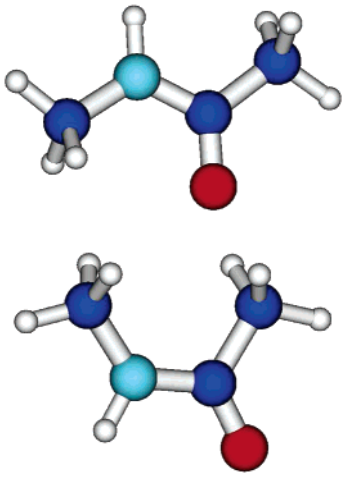

Figure 1. Schematic representation of NMA in the trans (top) and cis (bottom) conformation.

2. NMA as a Model Molecule

NMA (*N*-methylacetamide), a single amide "blocked" with methyl groups at both extremities, is the simplest building block model of the repeated peptide linkage of peptides and proteins. Figure 1 shows a schematic representation of the two conformers of the NMA molecule studied in this work, *trans*-NMA and *cis*-NMA. Structure and energetics of NMA have been the subject of numerous experimental and theoretical investigations.[18−26] Similarly the molecule has been a popular model system in infrared and Raman vibrational experiments either in the gas phase[7,27−29] or in solution.[7,30−34] Of particular concern in these studies were the possible relations between vibrational signatures of amide modes of this peptide model and those of secondary structures of peptides and proteins.

First theoretical studies were aimed at the understanding of electronic and geometrical properties of gas-phase NMA.[18−26] In these earlier studies, great emphasis was put on the electronic properties of the HNCO peptide linkage. Also the hydration structure of NMA has been investigated using computational methods, either ab initio geometry optimizations[19,20,35] with a few water molecules placed close to the main hydrophilic sites of NMA or applying classical molecular dynamics simulations of NMA surrounded by a few hydration shells. A variety of Hamiltonians were used, ranging from standard nonpolarizable force fields[36] to polarizable force fields[37] and mixed QM/MM dynamics.[23] Very recently, Martyna et al.[38] have studied the solvation structure of NMA using Car−Parrinello dynamics. As part of a careful classical molecular dynamics simulation of aqueous NMA involving a polarizable force field, Kato et al.[37] have performed a benchmark calculation of the infrared spectrum, which highlighted the importance of including polarization for an accurate description of the IR absorption. In a very recent work with ab initio geometry optimizations and harmonic vibrational analyses, Besley[35] has also shown the importance of incorporating a few layers of solvent around the solute (through a continuum model in that study) for a better description of the IR absorption patterns of aqueous NMA.

3. Computational Method

3.1. Ab Initio Molecular Dynamics. The DFT-based Car−Parrinello simulations performed in this work follow the general setup of previous ab initio molecular dynamics simulations. We used the Becke, Lee, Yang, and Parr (BLYP) gradient-corrected functional[39,40] for the exchange and correlation terms. 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[41] 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*), while H atoms are treated as an *s* potential with a 0.5 au radius. We should mention that the 70 Ry energy cutoff is mostly determined by the convergence of the H and O pseudopotentials. Energy expectations are calculated in reciprocal space using the Kleinman−Bylander transformation.[42]

Simulations were performed at constant volume using a fictitious electron mass of 500 au, a time step of 5 au (0.12 fs), and periodic boundary conditions. Gas-phase simulations were carried out in a cubic box of length 12 Å. This box size has been chosen after performing a series of wave function optimizations of NMA conformers in cubic boxes of increasing length. We found that, from 12 Å on, the electronic energy of the isolated molecule is converged within a criterion of 10^{-4} au, which ensures that periodic images of the molecule do not interact. Liquid simulations have been performed in a cubic box of length 12.07 Å, which is a compromise between spatial extent and duration of the MD runs. Starting configurations have been prepared using classical force field simulations, where we chose the precise number of water molecules in the simulation cell; this is done in order to fix the correct density of such a small simulated liquid system. We refer the reader to refs 12 and 43 for more details.

The cis/trans isomerization barrier of NMA in aqueous solution at room temperature is too high[38] to observe spontaneous isomerization on the molecular dynamics time scale. The infrared spectral features of *trans*-NMA and *cis*-NMA will thus be investigated by running two separate dynamics simulations each for a different conformation. In the *trans*-NMA model system the solute is surrounded by 50 water molecules and *cis*-NMA by 52 water molecules, which corresponds to two complete layers of solvent, and part of the third one, around the solute. These system dimensions are similar to the Car−Parrinello simulation of aqueous uracil performed previously.[12] The Car−Parrinello molecular dynamics simulations consist of two steps: An equilibration phase of 1 ps with a control of temperature through velocity rescaling, followed by data collection over trajectories of 9.35 and 6.35 ps, respectively, for aqueous *trans*-NMA and *cis*-NMA and 19.6 and 8 ps for gas-phase *trans*-NMA and *cis*-NMA. In that step, molecular dynamics are strictly microcanonical. The average ionic temperatures obtained along the simulations were 320 K (aqueous *trans*-NMA), 302 K (aqueous *cis*-NMA), 20 K (isolated *trans*-NMA), and 15 K (isolated *cis*-NMA). Hydrogen atoms were treated as classical particles with their true mass (1836 au). All simulations were carried out with version 3.7.1 of the CPMD ab initio molecular dynamics package.[44]

3.2. Computation of the Solute Spectrum. The first method applied here for the calculation of the infrared (IR) absorption coefficient $\alpha(\omega)$ makes use of the relation to the dipole time correlation function[45]

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

where $\beta = 1/kT$, $n(\omega)$ is the refractive index, c is the speed of light in a vacuum, V is the volume, and $\mathbf{M}$ is the total dipole moment of the system, which is the sum of the ionic and electronic contribution.[15] The angular brackets indicate a statistical average. In this formula, we have taken into account a quantum correction factor (multiplying the classical line shape) of the form $\beta\hbar\omega/(1 - \exp(-\beta\hbar\omega))$, which was shown to give the most accurate results on calculated IR amplitudes.[12] For a complete discussion on quantum corrections, we refer the reader to ref 77. The IR spectra in this work are given as products $\alpha(\omega)n(\omega)$ expressed in cm^{-1} (decadic linear absorption coefficient) as a function of reciprocal wavenumber ω in cm^{-1}. The spectra have been smoothed with a window filtering applied in the time domain, i.e., each term of the correlation function $C(t)$ is multiplied by a Gaussian function $\exp(-0.5\sigma(t/t_{\max})^2)$, where $t_{\max}$ is the length of the simulation, and σ is 10 for gas-phase simulations and 40 for liquid simulations. The periodic nature of an MD cell is a fundamental complication for the determination of the total dipole moment $\mathbf{M}$. This problem has been resolved by the modern theory of polarization[46,47] by describing polarization in terms of a Berry phase. This methodology was developed by Silvestrelli and Parrinello[14−16] in the context of Car−Parrinello molecular dynamics simulations.

When interested in the infrared spectrum of a solute immersed in liquid water, we face additional difficulties, namely separation of the absorption in solute and solvent contributions. Due to the limited statistics the subtraction method used in experiment is not an option for simulation. The approach applied previously in our calculation of the IR spectrum of aqueous uracil[12] is based on a decomposition of charge density using the maximally localized Wannier Functions scheme of Marzari and Vanderbilt.[48] This enabled us to write the total electronic dipole moment $\mathbf{M} = \mathbf{m}_{\mathrm{NMA}} + \sum_{\mathrm{W}}\mathbf{m}_{\mathrm{W}}$ as the sum of the solute dipole moment and the sum of the solvent dipole moments. As a result, the total dipole correlation function is resolved in three contributions $C_{\mathrm{MM}}(t) = C_{\mathrm{W/W}}(t) + C_{\mathrm{W/NMA}}(t) + C_{\mathrm{NMA/NMA}}(t)$. The IR spectrum of solute NMA is calculated from only the Fourier transform of the self-correlation function $C_{\mathrm{NMA/NMA}}(t)$. As discussed in ref 12, one disadvantage of this decomposition is that coupling with the surrounding solvent molecules is taken into account only through intermolecular polarization. The cross term $C_{\mathrm{W/NMA}}(t)$, which includes the remaining IR contributions arising from correlations between the motion of solute and solvent, is ignored. In fact, this term is hard to compute. Opposite positive and negative contributions make it difficult to reach convergence within the short simulation

times of tens of picoseconds. Moreover while fully consistent in the sense that the molecular dipole moments rigorously sum to the cell dipole moment, the Wannier function method suffers from a certain arbitrariness in the definition of molecular dipole moment in solution.

An alternative approach to achieve a similar, but not necessarily equivalent, decomposition of the absorption signal is to return to the more fundamental current representation, involving the autocorrelation function of the total current, $\mathbf{j} = \dot{\mathbf{M}} = \mathrm{d}\mathbf{M}/\mathrm{d}t$:

$$n(\omega)\alpha(\omega) = \frac{2\pi}{3Vk_BT}\int_{-\infty}^{\infty}\langle \mathbf{j}(0)\mathbf{j}(t)\rangle e^{i\omega t}\mathrm{d}t \qquad (2)$$

Note again that in this formula we have taken into account a quantum correction factor of the form $\beta\hbar\omega/(1 - \exp(-\beta\hbar\omega))$. Since in the Born−Oppenheimer approximation the dipole moment of the system depends only on the atomic positions, one can make use of a chain rule to compute the current $\mathbf{j}(t)$ at time t

$$j^{\beta}(t) = \sum_{i,\alpha}\frac{\partial M^{\beta}}{\partial q_i^{\alpha}}(t)\frac{\mathrm{d}q_i^{\alpha}}{\mathrm{d}t}(t) = \sum_{i,\alpha}\frac{\partial M^{\beta}}{\partial q_i^{\alpha}}(t)v_i^{\alpha}(t) \qquad (3)$$

where $\mathbf{M}$ is the total dipole moment of the system, $\mathbf{q}_i$ is the position of atom i, and $\partial M^{\beta}/\partial q_i^{\alpha}$, $\alpha,\beta = x, y, z$, represent the components of the atomic polar tensor (APT) of atom i. The APT of atom i is thus defined as the derivatives of the dipole moment of the system with respect to infinitesimal displacements of atom i.[49,50] This tensor is directly related to intensities of infrared bands in the double harmonic approximation.[51] Numerous calculations of APT have been performed in the solid phase or in the gas phase.[49,50,52,53] Recently, a calculation of the APT of water molecules in liquid water has been performed.[54] For our system, the APT tensor was calculated for each atom in the solution using Density Functional Response Theory.[50] This was repeated every 1.2 fs over a ≈1 ps interval taken from the full trajectory of *trans*-NMA in solution. We have used the Putrino et al.[55] implementation of the linear response to an applied electric field in the CPMD code.[44] In this scheme, the macroscopic polarization of the periodically replicated cell is defined using the Berry phase approach of Resta.[56,57] The APT tensor for the whole system is then obtained from the Maxwell relation

$$\frac{\partial M^{\beta}}{\partial q_i^{\alpha}} = \frac{\partial F_i^{\alpha}}{\partial \epsilon^{\beta}} = \frac{\partial^2 E_{\mathrm{tot}}}{\partial \epsilon^{\beta}\partial q_i^{\alpha}} \qquad (4)$$

where ϵ is an applied uniform electric field and $\mathbf{F_i}$ is the force acting on particle i. This allows for the calculation of the $3N$ elements of the system APT tensor through only three linear response calculations ($\beta = x, y, z$).[50] Using the same methodology, we have computed the APT of *trans*-NMA in the gas phase at the equilibrium configuration. Relation of APT tensors to infrared intensities is well-known from quantum chemistry.[78−80]

To resolve the infrared spectrum of the whole system into solvent, solute and cross contributions, we separated the total current of the system into a solvent and a solute term

$$j^{\beta}(t) = j^{\beta}_{\mathrm{water}}(t) + j^{\beta}_{\mathrm{NMA}}(t) \qquad (5)$$

by restricting the sum in eq 3 to atoms of the solvent or the solute only. This natural decomposition of the current then provides a decomposition of the infrared spectrum which is different from the one obtained through the maximally localized Wannier orbitals. The infrared spectrum of NMA is now defined as

$$n(\omega)\alpha_{\mathrm{NMA}}(\omega) = \frac{2\pi}{3Vk_BT}\int_{-\infty}^{+\infty}\langle \mathbf{j}_{\mathrm{NMA}}(0)\mathbf{j}_{\mathrm{NMA}}(t)\rangle e^{i\omega t}\mathrm{d}t \qquad (6)$$

A major argument in favor of the current scheme is that any partitioning of the charge density is avoided. We can also take advantage of the calculation of the APT tensor of each atom of NMA in the gas phase and in the liquid phase to get the transverse effective charges, or Born charges, defined as

$$Z_i^B = \frac{1}{3}Tr\frac{\partial M^{\beta}}{\partial q_i^{\alpha}} \qquad (7)$$

which corresponds to the isotropic component of the APT of atom i. In the gas phase, this charge was shown to correlate very well with other charge definitions,[52] although the actual value was found to be rather sensitive to the approximation used for the density functional. We will discuss later the values of these charges for *trans*-NMA as well as their variations when going from the gas phase to the liquid.

4. Structural Analysis

4.1. Intramolecular Properties. Table 1 reports values of the main internal geometrical distances and angles of the *trans*-NMA and *cis*-NMA conformers obtained by averaging over the ab initio molecular dynamics trajectories, either in the gas or aqueous phase. Atoms are labeled according to the following convention (for molecular geometries see Figure 1): N is the amide nitrogen atom, C is the carbon of the carbonyl group, and C_C and C_N denote the two carbon atoms of the methyl groups located at the C-terminal and N-terminal part of the molecule. In the gas phase, it can be seen that bond lengths are very similar for both conformers. The largest difference is observed for the amide N−H bond which is increased by 0.003 Å in *cis*-NMA. Backbone distances, i.e., N−C, N−CH_3, and C−CH_3, are almost equal in both conformers, and we do not see any geometrical differences in C−H bonds whether the methyl group is located at the N-terminal or the C-terminal part of the molecule. In *cis*-NMA, the average distance between the carbonyl oxygen and amide hydrogen, i.e., 2.432 Å, is slightly too long to be considered as a proper intramolecular hydrogen bond.

Values of intramolecular distances obtained in our gas-phase ab initio molecular dynamics at low temperature are very similar to those obtained with high level ab initio geometry optimizations.[21,24,25] Depending on the ab initio method (HF, DFT, MP2) and on the basis set used,[19,21,24,25] the calculated C═O distance of *trans*-NMA lies between

Table 1. Geometrical Distances and Angles of *trans*-NMA and *cis*-NMA Averaged over the Car−Parrinello Molecular Dynamics Trajectory

	C═O	N−H	O···H	N−C	O−C−N−H	C−C−N−C
trans-NMA[a]	1.238 ± 0.006	1.018 ± 0.007	3.183 ± 0.011	1.376 ± 0.006	179.93 ± 4.62	180.05 ± 4.15
cis-NMA[a]	1.239 ± 0.005	1.021 ± 0.007	2.432 ± 0.017	1.378 ± 0.006	0.718 ± 3.351	−1.266 ± 5.845
trans-NMA[b]	1.268 ± 0.022	1.031 ± 0.032	3.186 ± 0.053	1.351 ± 0.028	173.98 ± 10.75	178.46 ± 10.70
cis-NMA[b]	1.272 ± 0.021	1.032 ± 0.023	2.423 ± 0.099	1.350 ± 0.024	−3.330 ± 10.540	−0.204 ± 12.896

	N−CH_3	C−CH_3	C−H (N-terminal)	C−H (C-terminal)
trans-NMA[a]	1.469 ± 0.006	1.529 ± 0.008	1.097 ± 0.003	1.098 ± 0.003
cis-NMA[a]	1.470 ± 0.005	1.527 ± 0.011	1.098 ± 0.003	1.097 ± 0.004
trans-NMA[b]	1.477 ± 0.031	1.520 ± 0.035	1.101 ± 0.021	1.098 ± 0.016
cis-NMA[b]	1.482 ± 0.039	1.514 ± 0.031	1.100 ± 0.023	1.100 ± 0.021

[a] Isolated molecule at ∼20 K. [b] Hydrated molecule at ∼300 K. Distances are given in angstroms, and angles in degrees.

1.221 and 1.237 Å, the N−C distance is comprised in the 1.364−1.383 Å interval, and the amide N−H distance lies in the 1.008−1.016 Å interval. The average distances obtained in our molecular dynamics of *trans*-NMA and *cis*-NMA are virtually equal to the ones calculated by DFT/BLYP geometry optimizations in ref 21, except for the N−C distance which appears slightly shorter (− ∼0.010 Å) in our dynamics. As usual with DFT calculations, it is found that distances obtained with the BLYP functional are overestimated compared to the B3LYP functional.[21] Comparing our results to experiment[58,59] for *trans*-NMA, we find that the C═O distance is overestimated by ∼0.010 Å, while the N−C distance is underestimated by ∼0.010 Å.

As can be anticipated, we find that C═O and N−H intramolecular distances are increased (almost equally by ∼0.030 Å) in aqueous solution, due to the formation of intermolecular hydrogen bonds between these groups and the surrounding water molecules. In aqueous solution, the backbone N−C distance is very much shortened (by ∼0.025 Å) for both NMA conformers. This is also observed for the backbone C−CH_3 distance in the trans (−0.009 Å) and cis (−0.013 Å) conformers. In contrast, the N−CH_3 bond length is increased from the gas to the liquid phase (+0.008 Å for *trans*-NMA and +0.012 Å for *cis*-NMA). Thus, in aqueous solution, we observe that the N−C and C−CH_3 backbone bonds are contracted, whereas the N−CH_3 bond is extended. As anticipated for methyl hydrophobic groups, C−H bond distances are not affected by the presence of solvent. The modifications of intramolecular distances observed here are consistent with previous geometry optimization calculations[20,26] with a few water molecules hydrogen-bonded to the N−H and C═O groups. The presence of two layers of solvent around NMA simulated in our work amplifies the variations observed with only a few solvent molecules. Overall, due to the higher temperature in the liquid simulation, the fluctuations around the mean distance or mean angle values displayed in Table 1 are higher in the liquid phase than in the gas phase.

In Table 1 we have also reported the average values of the OCNH and CCNC peptide bond dihedral angles obtained over the whole MD trajectory. Each NMA conformer did remain in its initial trans- or cis-orientation (OCNH angle), be it in the gas phase or in solution. We recall that gas-phase calculations have shown an energy difference in the range ∼2−3 kcal/mol between the trans and cis conformers.[21,23] According to the calculations of Martyna et al.[38] solvation enhances this barrier by ∼2−3 kcal/mol. We note that the OCNH dihedral angle of gas-phase *trans*-NMA obtained in our work is identical to that in ref 21, while we predict an angle of 0 degree for *cis*-NMA instead of 6.05 degrees.[21] We finally observe that the mean values of the OCNH dihedral angles of *trans*- and *cis*-NMA are decreased by a few degrees when going from the gas phase to the solution.

4.2. Radial Pair Correlation Functions. As usual in simulations of liquids, determination of radial distribution functions (hereafter denoted rdfs) can provide a good first impression of hydrogen bonding among the solvent molecules and between the solute hydrophilic groups and the solvent. The relevant rdfs are reported in Figure 2, respectively, for NMA amide hydrogen/water oxygens (H−Ow, top left), NMA carbonyl oxygen/water hydrogens (O−Hw, top right), NMA methyl carbons/water oxygens (CH_3−Ow, bottom left), and water/water (Ow−Hw, Ow−Ow, and Hw−Hw, bottom right of Figure 2). In each figure, we compare results obtained for aqueous *trans*-NMA (plain lines) to those obtained for aqueous *cis*-NMA (dashed lines).

The broad picture of NMA hydration emerging from the Car−Parrinello molecular dynamics simulations follows what can be expected for this kind of organic molecule. H−Ow and O−Hw radial distribution functions display the general features of hydrogen bonding between hydrophilic sites, i.e., between N−H amide and water oxygens or between C═O carbonyl and water hydrogens. The first peak of the H−Ow rdfs is located at 2.0 and 1.9 Å for *trans*-NMA and *cis*-NMA, respectively, while the first minimum is found at 2.6 and 2.5 Å. The first peak of O−Hw rdfs appears at slightly shorter distances, 1.8 and 1.7 Å, whereas the position of the first minimum is at 2.5 Å for both conformers. These features are entirely consistent with the formation of strong hydrogen bonds between amide groups and water molecules, on the one hand, and between carbonyl groups and water molecules on the other hand. We note that H−Ow and O−Hw separations as measured by first maximum in rdfs are always 0.1 Å shorter for aqueous *cis*-NMA. By integrating rdfs up to the position of the first minimum, we find that the number of water molecules hydrogen bonded to the amide group is 0.87 for *trans*-NMA and 0.91 for *cis*-NMA, while for the

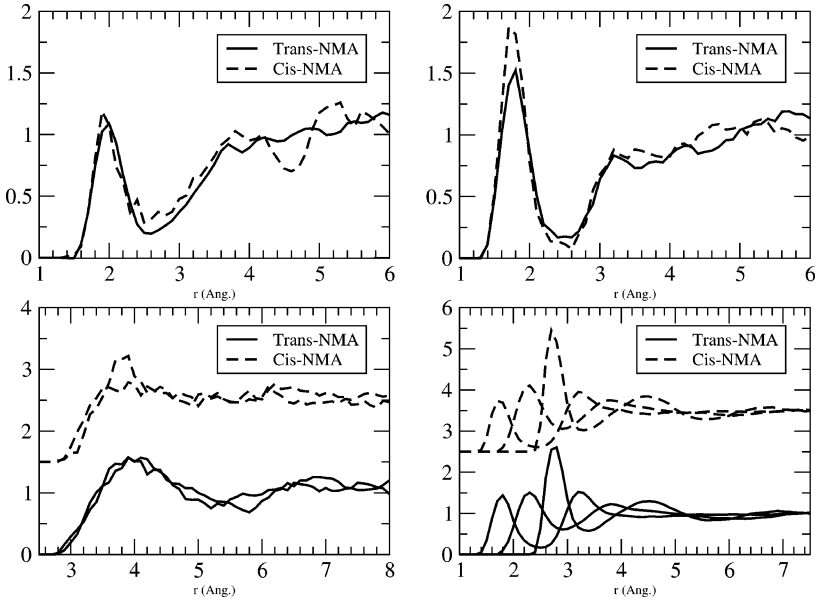

Figure 2. Solute−solvent and solvent−solvent radial distribution functions (RDF) calculated from Car−Parrinello molecular dynamics simulations of aqueous NMA. Conventions: H is NMA amide hydrogen atom, O is NMA carbonyl oxygen atom, C is one of the two NMA carbonyl carbon atoms, Ow and Hw are respectively the oxygen and hydrogen atoms of water molecules. RDF reported here: H−Ow (top left), O−Hw (top right), C−Ow (bottom left, with both carbonyl represented), Ow−Ow, Ow−Hw, Hw−Hw (bottom right). RDFs for *trans*-NMA are always in solid lines, and those for *cis*-NMA are in dashed lines. Distances are in angstroms.

carbonyl group the number is 2.0 for *trans*-NMA and 2.22 for *cis*-NMA. The total number of water molecules coordinated with the hydrophilic amide and carbonyl groups is thus slightly higher for *cis*-NMA than for *trans*-NMA. This results in a higher amplitude of the first O−Hw peak for *cis*-NMA (top right of Figure 2), whereas the amplitude of the first minimum is identical for both O−Hw rdfs. This difference suggests that water molecules are more tightly H-bonded to the C═O group of *cis*-NMA than to that of *trans*-NMA.

We observe a well-defined second maximum in the H−Ow rdf of aqueous *cis*-NMA, located around 4 Å with a finite width of ∼1 Å. Taking into account the characteristic distances displayed in our rdfs, this suggests that some water molecules belonging to the second hydration shell around the amide hydrogen of *cis*-NMA are also part of the first hydration shell of the carbonyl oxygen atom. This type of bridge formation is only possible for a cis conformer where the N−H and C═O are placed close to each other. This also implies that water molecules H-bonded to the amide group of *cis*-NMA are separated from the ones H-bonded to the carbonyl group, i.e., there are no water molecules shared by both groups in their first hydration shell. This situation is different from what would be expected in the geometry optimization (at 0 K) of *cis*-NMA with one water molecule, where the water molecule would be found to make a bridge so as to form a stable six-membered ring including the CO, CN, and NH bonds of *cis*-NMA. At this point, it should be reminded that the finite temperature (300 K) and the presence of a bulk solvent in our simulation lead to a completely different view. In particular, as will be shown below (and already found in our previous investigation[61]), solvent−solvent H-bonds are always energetically more favorable than solute−solvent H-bonds. At 300 K, water molecules bind to the solute, while maintaining H-bonds with the bulk water molecules so as to stabilize the collective water−water H-bond network.

First peak positions obtained in our ab initio dynamics of aqueous *trans*-NMA simulation are similar to the results of another Car−Parrinello MD study[38] and also of QM/MM calculations[23] as well as classical dynamics with a polarizable force field[37] or a QM-RISM calculation.[60] There are, however, minor discrepancies. QM/MM calculations give a first peak position for O−Hw at a slightly larger distance than in our work, i.e., 2.10 and 1.95 Å depending on the ab initio level. Furthermore, QM-RISM and classical O−Hw radial distribution functions of *trans*-NMA display a marked narrow second peak of low amplitude, located between 3.0 and 4.0 Å. This finding appears also in the QM/MM calculations at the Hartree−Fock level and a rather poor basis set. This feature is absent in our ab initio MD.

The main differences between our work and the previous studies come from the strength that can be attributed to the intermolecular H-bonds formed between solute and solvent. A rough estimate is given by the ratio of the amplitude of the first peak (denoted g^{max}) to the amplitude of the first minimum (denoted g^{min}) of the radial distribution functions. Using our data, this gives a H−Ow H-bond strength for aqueous *trans*-NMA of $g^{max}/g^{min} \sim 5$, which is a bit stronger compared to the $g^{max}/g^{min} \sim 4$ obtained using a polarizable force field and ∼3.6 with QM/MM (again with variations depending on the ab initio level of the QM part). The ab initio MD bond strength appears much stronger than in the QM-RISM calculation where a value of ∼2.0 is obtained. In contrast, our O−Hw H-bond strength is lower than the ones obtained in using QM/MM and MM methods ($g^{max}/g^{min} \sim 7.5$ versus ∼8.5 in QM/MM and ∼9.5 with polarizable

Table 2. Geometrical Distances between NMA Atoms and Wannier Centers Obtained by Averaging over the CPMD Trajectory of *trans*-NMA and *cis*-NMA

	trans-NMA[a]	*trans*-NMA[b]	*cis*-NMA[a]	*cis*-NMA[b]
N−WC N−C bond	0.475 ± 0.003	0.495 ± 0.032	0.459 ± 0.039	0.496 ± 0.033
N−WC N−C bond	0.476 ± 0.003	0.509 ± 0.030	0.480 ± 0.038	0.506 ± 0.031
N−WC N−H bond	0.607 ± 0.003	0.586 ± 0.013	0.602 ± 0.003	0.585 ± 0.011
N−WC N−CH3 bond	0.627 ± 0.002	0.626 ± 0.010	0.634 ± 0.002	0.626 ± 0.013
O−WC lone pair	0.314 ± 0.002	0.338 ± 0.014	0.318 ± 0.001	0.334 ± 0.011
O−WC lone pair	0.308 ± 0.002	0.323 ± 0.014	0.308 ± 0.001	0.342 ± 0.013
O−WC C═O bond	0.499 ± 0.002	0.469 ± 0.013	0.502 ± 0.002	0.463 ± 0.016
O−WC C═O bond	0.499 ± 0.002	0.463 ± 0.015	0.502 ± 0.002	0.467 ± 0.012
C_C−WC C_C−C bond	0.756 ± 0.004	0.764 ± 0.019	0.756 ± 0.006	0.762 ± 0.018
C_C−WC C_C−H bond	0.734 ± 0.009	0.721 ± 0.025	0.728 ± 0.004	0.718 ± 0.018
C_C−WC	0.726 ± 0.009	0.720 ± 0.021	0.722 ± 0.002	0.717 ± 0.020
C_C−WC	0.722 ± 0.003	0.721 ± 0.021	0.728 ± 0.003	0.718 ± 0.022
C_N−WC C_N−H bond	0.731 ± 0.009	0.721 ± 0.021	0.728 ± 0.003	0.727 ± 0.021
C_N−WC	0.728 ± 0.010	0.721 ± 0.021	0.722 ± 0.002	0.726 ± 0.016
C_N−WC	0.733 ± 0.008	0.721 ± 0.021	0.728 ± 0.003	0.728 ± 0.021

[a] Isolated molecule at ∼20 K. [b] Hydrated molecule at ∼300 K. Distances are given in angstroms.

force fields) but still well above the ∼6.4 determined in the QM-RISM work. Finally, we draw attention to the very good agreement between our results and Martyna's,[38] who used almost identical AIMD methods but a smaller model system consisting of 27 water molecules.

Also very instructive is the comparison of the H−Ow and O−Hw H-bond strength for *trans*-NMA and *cis*-NMA as implied by the structure of the ab initio MD rdfs. The H−Ow H-bonding is comparable in both systems (same values of g^{max} and g^{min}), whereas the O−Hw H-bond is stronger in *cis*-NMA ($g^{max}/g^{min} \sim 9$ versus ∼7.5). As previously emphasized, this can be traced back to water molecules more tightly H-binding to the carbonyl group of *cis*-NMA than to *trans*-NMA. This is also connected to the presence of a well-defined second peak in the H−Ow rdf of aqueous *cis*-NMA. Turning to the nonpolar sites in the molecule, Figure 2 (bottom left) shows all the rdfs involving the carbon atom of the CH_3 methyl groups. There is no indication of formation of hydrogen bonds between these groups and surrounding water molecules (first peaks located around 4.0 Å), and we do not see any distinction between C- or N-terminal methyls. The slightly higher amplitude of the peak observed for the N-terminal CH_3 of *cis*-NMA could be related to the more ordered second hydration shell around the N−H group as we have described previously.

Radial distribution functions among the solvent water molecules (Hw−Hw, Ow−Ow, and Hw−Ow) are reported in Figure 2 (bottom right). We find that each water molecule has 4.1 and 4.0 water neighbors on average for *cis*- and *trans*-NMA, which is similar, if perhaps somewhat smaller than the values in pure liquid water (≈4.3). Comparison of H-bond strengths estimated by the ratio g^{max}/g^{min} for all the computed rdfs (solute−solvent and solvent−solvent) leads to the following order: Hw−Ow > O−Hw > H−Ow, for both *trans*- or *cis*-NMA. Thus, H-bonds among the solvent molecules are always more beneficial than solute−solvent H-bonds. This result was already observed in the ab initio molecular dynamics of aqueous uracil.[61]

A comparison of the solvent−solvent rdfs of Figure 2 to pair distributions for the pure liquid can give a good indication of the extent to which the solvent is disturbed by the presence of the solute. The question is particularly relevant in view of our limited system dimensions. The rdfs of Figure 2 in fact appear very similar to the rdfs of (BLYP) homogeneous liquid water of refs 62−64 which, moreover, are in fair agreement with experiment. This suggests that solvent in our system essentially retains the structure of the pure liquid. However, the results of refs 62−64 have recently been questioned.[65] The rapid advances in computer power of recent years have made it possible to extend the length of the trajectories routinely to 20−30 ps and in exceptional cases even to longer times. Similarly, current AIMD simulations are carried using computationally more demanding implementations of the ab initio MD algorithm which approach the ideal Born−Oppenheimer limit closer than in the original simulations, using either smaller values of fictitious electron mass[65] or Born−Oppenheimer dynamics.[66] These new studies show the rdfs of BLYP water is significantly more structured than suggested by the first generation of MD simulations.[65] Even if the issue is far from settled,[81] the consensus seems now that a converged BLYP result at ambient conditions actually corresponds to the experimental structure of undercooled water at a substantially lower temperature (−20 °C).[66] The way this is manifested in practice is that the motion of the water (diffusion) slows down as the run progresses. Such a relaxation process was not observed in the present simulation of the model NMA solution. Note that the value of the fictitious electron mass of 500 au we have applied is significantly lower than the 800 au used in refs 62−64. Our trajectories, however, are still relatively short (∼10 ps). While it cannot be ruled out that further relaxation will take place on a longer time scale, the stability of our system for the duration of the simulation suggests that the system has settled in a (possibly) metastable state with a solvent structure and mobility in fair agreement with bulk liquid water.

5. Electronic Analysis

5.1. Wannier Functions and Born Charges. Table 2 reports the average positions of the Wannier centers (WC) obtained

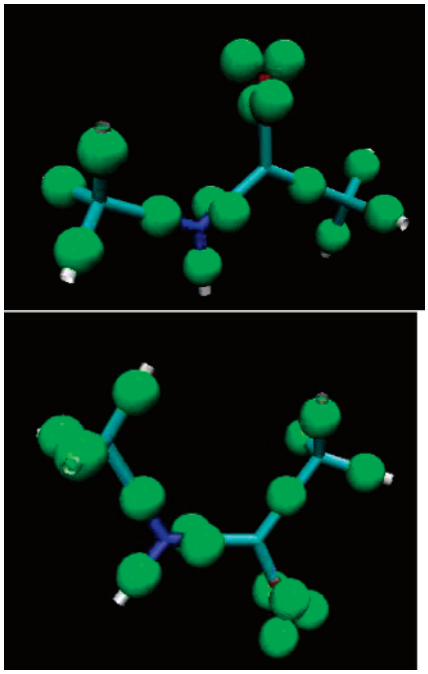

Figure 3. Representation of the positions of the Wannier centers (spheres) as calculated for *trans*-NMA (top) and *cis*-NMA (bottom). The Wannier center coordinates are given in Table 2 for the gas and aqueous phase.

from our gas- and aqueous-phase CPMD trajectories of *trans*- and *cis*-NMA. Wannier centers are represented in Figure 3 either by spheres centered on or near covalent bonds or by spheres that correspond to lone pairs of oxygen and nitrogen atoms. The same trends are observed for *trans*- and *cis*-NMA upon going from the gas to the aqueous phase. First, electrons along the N−H amide bond can be seen to be pulled toward the nitrogen atom in the liquid phase. This is expected from the formation of a hydrogen bond between N−H and surrounding water molecules. Similarly, we observe that the oxygen lone pairs are pushed out, also as a result of H-bonding, while electrons of the double bond C═O are displaced toward the oxygen atom. The polarization of the N−H and C═O bonds induced by the hydrogen bonds, is accompanied by a displacement of the electrons along the N−C bond away from the nitrogen atom, as indicated by the increase in nitrogen-Wannier centers distances (along that bond) in response to solvation. The electrons move closer to the middle of the bond suggesting that the N−C is stronger in the liquid, which is also consistent with the observed shrinking of the N−C bond length. This is a consequence of the nature of the peptide bond. Finally, we see that electrons along the $C-CH_3$ skeleton bond are pulled toward the carbonyl carbon in the liquid phase.

Born charges calculated according to eq 7 for the atoms of *trans*-NMA provide supplementary data to quantify and analyze the rearrangement of electrons, when going from the gas phase to the liquid phase. The results of this computation are listed in Table 3. Summing the charges over all atoms of the *trans*-NMA molecule, we verified that, as required by the sum rule for Born charges, the molecule is strictly charge neutral in the gas phase (absolute total charge less than 7×10^{-4} *e* due to numerical uncertainties). In

Table 3. Born Charges Calculated along the Trajectory of Gas-Phase *trans*-NMA[a] and Aqueous-Phase *trans*-NMA[b] (Units of Electrons)

atom type	gas phase[a]	aqueous phase[b]
C	1.049966	1.273998
O	−0.784995	−1.222748
N	−0.697676	−0.869836
H (amide)	0.161829	0.398282
C_C	−0.057427	−0.092539
C_N	0.339080	0.310198
H (C_C)	−0.011471	0.029868
H (C_C)	0.018505	0.036330
H (C_C)	0.020947	0.046038
H (C_N)	−0.003725	0.022462
H (C_N)	−0.039359	0.026497
H (C_N)	0.003645	0.018143

solution, however, *trans*-NMA acquires an average charge of −0.025 *e*. This relative small excess charge most likely originates in the dynamical contribution to the Born charge and should not be interpreted as a solvent to solute charge transfer. The dynamical charge measuring the response of the electronic charge density to displacements of the nuclei can lead to sometimes counterintuitive estimates of total charge. In particular in situations where the charge cloud lags behind, increasing the value of the dipole moment, the dynamical charge enhances the static contribution (for example the Born charge of a Na^+ ion in crystalline NaCl is +1.1 *e*).

Solvent induced changes of individual Born charges support the above view of electron flow based on the WC analysis. The C═O oxygen is made more negative by 0.44 *e*. The positive charge of the carbonyl carbon atom is increased by 0.22 *e*. On the N−H side, we observe that the hydrogen atom becomes more positive by 0.24 *e*, while the negative nitrogen atom is decreased by another 0.17 *e* in the aqueous phase compared to the gas phase. Alongside this polarization of the C═O and N−H bonds, most certainly arising from H-bonding with surrounding water molecules, it can be seen that the C═O end (including the C_CH_3 group) of *trans*-NMA gains 0.16 *e* in total which comes from the N−H group of the molecule. This charge transfer is again related to the special character of the peptide bond. Compared to atoms of the HNCO amide system, hydrogen atoms are globally chargeless, either in the gas phase or in the liquid phase.

The above picture can be rationalized with the help of the resonance model for the peptide bond: $C(=O)-NH \leftrightarrow C(-O^-)=NH^+$. In the gas phase, the neutral form is the most stable,[67] while in the aqueous phase we would expect the hydrogen bonding with the neighboring water molecules to stabilize the zwitterionic form with a more negatively charged oxygen (playing the role of an electron donor for the H-bond) and a positively charged N−H group (more ready to accept electrons from surrounding waters). This in turn is consistent with the observed charge transfer from the N−H end to the C═O end of *trans*-NMA (indicated by variations of the Born charges and WC displacements) and a strengthening of the N−C bond (shorter in the aqueous phase) and a weakening

Table 4

MD[a]	assignments[b]	expt[c]	expt[d]
3344	ν(N−H)		
2918, 2903, 2861	ν(C−H)		
1609	ν(C═O), δ(N−H) (amide I)	1707	1731, 1714
1458	δ(N-H), ν(N−C) (amide II)	1511	1499
1430	$\delta(C-H)_N$, ν(N−C)	1472	
1407	$\delta(C-H)_C$, ν(N−C)	1446	
1384	$\delta(C-H)_C$	1432	
1367	$\delta(C-H)_N$	1419	
1327	$\nu(C-CH_3)$, $\delta(C-H)_C$	1370	∼1420
1189	ν(N−C), δ(N−H) (amide III) + $\nu(C-CH_3)$, $\nu(N-CH_3)$	1266	1255
1098	$\nu(N-CH_3)$, $\delta(C-H)_N$	1168	
1015, 1005	ν(N−C), $\nu(N-CH_3)$, $\nu(C-CH_3)$, δ(N−H), CH_3 rockings	1089	

[a] Band positions (cm^{-1}) in the vibrational density of states of isolated *trans*-NMA computed at 20 K. [b] Assignments according to the method described in the Supporting Information. The notation of the bands are described in the text. [c] Experimental IR spectrum of *trans*-NMA in an argon matrix at 20 K.[27] [d] Experimental IR spectrum of *trans*-NMA in the gas phase at ∼100 K.[7]

of the C═O bond (longer in the liquid phase). This was already suggested by previous studies by Torii et al. for the lowest excited state of *trans*-NMA which has a clear zwitterionic form in the gas phase and, reversely to the ground state, a less pronounced zwitterionic form in aqueous phase.[67] The important consequences of these charge transfers on the IR intensities in solution will be examined in a later section.

5.2. Dipole Moments. Averaging the dipole moments of isolated *trans*-NMA and *cis*-NMA over the low temperature trajectories gives 3.99 D and 4.38 D, respectively. The dipole moment of *cis*-NMA is thus higher than the one of *trans*-NMA, an ordering also obtained in other theoretical works[23,24,60] using geometry optimizations of *trans*- and *cis*-NMA at various levels of calculations. The values obtained here are also very similar to those of refs 23, 24, and 60. We note that our 3.99 D value for *trans*-NMA is very close to the experimental value of 3.85 D measured in benzene.[68,69] In solution, the intramolecular charge transfers described in the previous section lead to a substantial increase of the dipole moments of both isomers. Adding the positions of the nuclei and centers of the maximally localized Wannier functions discussed in section 5.1 we find that solvent polarization enhances the average dipole moments to 6.96 D for *trans*-NMA and 7.33 D for *cis*-NMA. *cis*-NMA remains therefore the more polar conformer in solution. Also the dipole difference is, on average, conserved, i.e., 0.39 D in the gas phase versus 0.37 D in aqueous solution. Our assignment of the higher dipole in solution to *cis*-NMA is in agreement with the QM-RISM work of Wei et al.,[60] whereas the reverse ordering was found in a PCM-B3LYP model using the SCRF theory.[24] The absolute values of the aqueous dipole moments predicted by both calculations are considerably lower than the ab initio MD estimates. The values of ref 24 are especially low, almost equal to the values that we obtain in the gas phase. Our dipole moments are also ∼1.2−1.4 D higher compared to the ones computed in ref 60.

6. Vibrational Density of States

To understand our calculated infrared spectra in terms of vibrationally active infrared modes, we have computed the vibrational density of states (VDOS) of the isolated and aqueous *trans*- and *cis*-NMA molecules. The VDOS is obtained by Fourier transformation of the velocity autocorrelation function of the atoms. In addition we have decomposed the VDOS according to each atom type in order to get an interpretation of the vibrational bands in terms of individual atomic motions. Regarding the 1000−2000 cm^{-1} frequency domain which is of particular interest in experiments on peptides and proteins, and specifically related to the amide stretching and bending modes used for the characterization of helices and sheets structures, this simple decomposition proves sufficient to properly assign each vibrational peak of NMA to specific atomic motions. A complication is that the rocking motions of methyl groups have also a spectral signature in the same spectral domain. We have thus calculated also the Fourier transform of the correlation function of the three principal axis components of the methyl groups, to get more specific spectral signatures of rocking motions. The details of all this analysis are given in the Supporting Information. For each species, the final assignments of the different peaks are summarized in Tables 4−7. The overall spectrum in the 1000−2000 cm^{-1} region can be discussed in terms of coupled stretching (ν) and bending (δ) motions, more specifically ν(C═O), ν(N−C), and δ(N−H) for the central peptide region and $\nu(N-CH_3)$, $\delta(C-H)_N$, $\nu(C-CH_3)$, and $\delta(C-H)_C$ for the N-terminal and C-terminal regions, respectively. The systematic coupling of ν(C═O) and δ(N−H) gives rise to the so-called amide I mode, and the one between ν(N−C) and δ(N−H) results in amide II and amide III. The $\nu(N-CH_3)$ and $\nu(C-CH_3)$ stretches appear always slightly coupled to the C−H bends of the attached methyl groups, $\delta(C-H)_N$ and $\delta(C-H)_C$. A collective backbone motion involving all the backbone stretches, the N−H bend, and the rocking of the methyl groups is always observable at the lowest frequency part of the spectrum.

6.1. Isolated NMA. The vibrational frequencies and the corresponding assignments of gas-phase *trans*- and cis-NMA are reported in Tables 4 and 6. It appears that the two spectra have very similar features in the 1000−1700 cm^{-1} spectral domain: one single band related to amide I around 1600 cm^{-1}, a broad band between 1300 and 1500 cm^{-1} that

Table 5

MD[a]	assignments[b]	expt[c]
3259, 3200		
1983, 2920, 2861		
1535, 1522	amide I	1617[c1], 1620[c2], 1625[c3]
1483, 1466	amide II	1577[c1], 1580[c2], 1582[c3]
1450−1325	$\delta(C-H)_N$, $\delta(C-H)_C$	1416[c1], 1418[c2], 1378[c1], 1378[c2]
1233	amide III + $\nu(N-CH_3)$, $\delta(C-H)_N$, $\delta(C-H)_C$	1314[c1], 1315[c2]
1098	$\nu(N-CH_3)$, $\delta(N-H)$, $\delta(C-H)_N$	1167[c2]
1013	$\nu(N-C)$, $\nu(N-CH_3)$, $\nu(C-CH_3)$, $\delta(N-H)$, CH_3 rockings	1097[c2]

[a] Band positions (cm^{-1}) in the vibrational density of states of aqueous *trans*-NMA computed at 320 K. [b] Assignments according to the method described in the Supporting Information. The notation of the bands are described in the text. [c] Experimental IR spectrum of aquous NMA from refs 31 (c1), 32 (c2), and 7 (c3).

Table 6

MD[a]	assignments[b]	expt[c]
3327	$\nu(N-H)$	
2950, 2920, 2853	$\nu(C-H)$	
1606	amide I	
1428	$\delta(C-H)_N$, $\delta(C-H)_C$	1485
1403	$\delta(C-H)_N$, $\delta(C-H)_C$, $\nu(N-C)$, $\nu(N-CH_3)$	1454
1369	amide II	1432
1331	$\nu(N-C)$, $\nu(C-CH_3)$, $\delta(C-H)_C$	1387
1259	amide III	1325
1115, 1085	$\nu(N-CH_3)$, $\delta(C-H)_N$	
1015, 1005	$\nu(N-C)$, $\nu(N-CH_3)$, $\nu(C-CH_3)$, $\delta(N-H)$, CH_3 rockings	1075

[a] Band positions (cm^{-1}) in the vibrational density of states of isolated *cis*-NMA computed at 15 K. [b] Assignments according to the method described in the Supporting Information. The notation of the bands are described in the text. [c] Experimental IR spectrum of *cis*-NMA in an argon matrix at high temperature.[27]

Table 7

MD[a]	assignments[b]
3255, 3187	
2890	
1560, 1534	amide I
1466, 1433	$\delta(C-H)_C$, $\delta(C-H)_N$
1407, 1382	amide II
1365, 1339, 1328	amide III + $\nu(C-CH_3)$, $\delta(C-H)_C$, $\delta(C-H)_N$
1127	$\nu(N-CH_3)$, $\delta(C-H)_N$
1009	$\nu(N-C)$, $\nu(N-CH_3)$, $\nu(C-CH_3)$, $\delta(N-H)$, CH_3 rockings

[a] Band positions (cm^{-1}) in the vibrational density of states of aqueous *cis*-NMA computed at 302 K. [b] Assignments according to the method described in the Supporting Information. The notation of the bands are described in the text.

contains the most intense contribution to the whole VDOS (∼1330 cm^{-1}), and two bands identically located in both spectra in the 1000−1200 cm^{-1} spectral region. The amide I mode is therefore insensitive to a change in conformation (1609 cm^{-1} for *trans*- versus 1606 cm^{-1} for *cis*-NMA). The main differences between the two vibrational spectra come from the position of the amide II and amide III modes. Amide II in *trans*-NMA, located at 1458 cm^{-1}, is red-shifted by 89 cm^{-1} in *cis*-NMA (1369 cm^{-1}). On the other hand, the amide III band which appears as a single isolated peak in the 1100−1300 cm^{-1} region is blue-shifted by 70 cm^{-1} from the cis to the trans form.

There are four peaks related to methyl C−H bending in *trans*-NMA (∼1430 cm^{-1}, ∼1407, 1384, and 1367 cm^{-1}) and only two for *cis*-NMA (1428 and 1403 cm^{-1}). The mean frequencies of the two bands above 1400 cm^{-1} are identical for *trans*- and *cis*-NMA (with slight variations in their assignments). We also note that while we can clearly distinguish vibrational contributions from the methyl group located at the N-terminal or C-terminal position of *trans*-NMA, this is not the case for *cis*-NMA for which the signals of both methyl groups merge. The two vibrational modes peaked at 1384 and 1367 cm^{-1} in the *trans*-NMA spectrum, related to δ(C−H) bending, have disappeared in *cis*-NMA. In this region, *cis*-NMA exhibits only one band (1369 cm^{-1}) but related to amide II vibration. The backbone $\nu(C-CH_3)$ stretch coupled to $\delta(C-H)_C$ bend (∼1330 cm^{-1}) is not affected by the trans/cis conformation. Finally, the 1000−1100 cm^{-1} vibrational region is identical for both conformers, though we note that the broad band at 1098 cm^{-1} in the *trans*-NMA spectrum, related to the backbone $\nu(N-CH_3)$ stretch and $\delta(C-H)_N$ bend, is split into two bands located at 1085 and 1115 cm^{-1} for *cis*-NMA.

Much attention has been given in the literature to vibrational calculations for *trans*-NMA in the gas phase using standard ab initio geometry optimizations and frequencies determined at the harmonic level.[19−22,25,27,70,71] Very little work has been devoted to *cis*-NMA.[21] Detailed vibrational assignments of gas-phase *trans*-NMA have been reported by Wang et al.,[70] Mayne et al.,[71] Hagler et al.,[21] and Gerber et al.[22] (who went beyond the harmonic approximation). One interesting point revealed by Gerber's study is a downshift of up to 40 cm^{-1} of anharmonic frequencies with respect to their harmonic approximation. In these recent studies, no frequency scaling has been applied, contrary to older ab initio calculations.[19,20,25,27] A fundamental understanding of the vibrational modes was originally given in Miyazawa's study et al.,[72] on which most of the subsequent ab initio studies rely. Overall, frequencies calculated in our work using the BLYP functional are downshifted (by up to ∼100 cm^{-1} for amide I) compared to frequencies calculated with hybrid functionals (e.g. B3LYP) or with wave function based ab initio calculations. 5−10% underestimation of frequencies is typical for the BLYP functional. Although our calculated vibrational band positions show a systematic red-shift compared to the work of Gerber et al.,[22] we find good agreement with Gerber's anharmonic calculation for the frequency separations between amide II and amide III and between amide III and the backbone stretches (respectively, 269/264 cm^{-1} and 174/138 cm^{-1}). Our amide I−amide II

frequency gap (151 cm^{-1}) is too low in comparison to the result of ref 22 (204 cm^{-1}). The vibrational bands assignments proposed from our molecular dynamics are very similar to refs 21, 22, 70, and 71, except for the methyl motions which, we think, are not accounted for properly within a harmonic approximation. We find that the C−H bending motions of the C-terminal methyl group of *trans*-NMA are packed in the same spectral domain (1407−1384 cm^{-1}), whereas they are split apart at the N-terminal end. Hagler et al.[21] and Gerber et al.[22] find a systematic alternation of these modes in the same spectral domain. We remark that Hagler et al. have used the original assignments of Ataka et al.,[27] for which some ambiguities remain. Finally, in agreement with Hagler's work, we find that the amide I and amide II modes of *cis*-NMA are separated by ∼240 cm^{-1}; on the other hand, our amide II−amide III separation is equal to 110 cm^{-1}, to be compared to 185 cm^{-1} in their work.

6.2. Aqueous NMA. Again, for the detailed analysis of the VDOS of each isomer in solution we refer to the Supporting Information. Tables 5 and 7 list the position and our interpretation of the vibrational bands. A very noticeable effect of the solvent on the VDOS is that the amide I stretching motion undergoes a red-shift (87 cm^{-1} for *trans*-NMA and 46−72 cm^{-1} for *cis*-NMA), whereas the amide II is blue-shifted (∼8−25 cm^{-1} and 3−38 cm^{-1}, respectively). The red-shift of amide I is consistent with the formation of hydrogen bonds between the carbonyl group and the surrounding water molecules, as was emphasized in the structural analysis of solute−solvent H-bonding and in the charge flow analysis of section 5. The red-shift computed for *trans*-NMA is about twice larger than obtained for aqueous uracil[12] in our previous study. This more pronounced red-shift could be the signature of very strong hydrogen bonds formed between the NMA carbonyl group and the solvent. Similarly the g^{max}/g^{min} ratios obtained in the present simulations of *trans*- and *cis*-NMA are much larger than the ones obtained for aqueous uracil[61] (∼3.7−4.5 versus ∼7.5−9.3 for NMA). The blue-shift of amide II is also the result of H-bonding between the amide group of NMA and the surrounding water molecules and was also observed in the vibrational spectrum of aqueous uracil.[12]

The comparison of the VDOS of aqueous *trans*-NMA and *cis*-NMA reveals also several interesting features. The two bands around 1000 and 1100 cm^{-1} are identically located in both spectra. The 1300−1500 cm^{-1} spectral domain displays similar peaks for *trans*- and *cis*-NMA, although we have seen that their interpretation leads to very different assignments. The most obvious differences are the δ(C−H) bending mode and the amide II mode, the positions of which are roughly inverted in both spectra. The amide II bands of *trans*- and *cis*-NMA are shifted by ∼80−100 cm^{-1} with respect to each other, just as was observed in the gas phase. Finally, the amide I mode is ∼30 cm^{-1} blue-shifted from *cis*- to *trans*-NMA, i.e., 1560−1534 cm^{-1} versus 1535−1522 cm^{-1}. We recall that this mode was peaked at ∼1606 cm^{-1} in both gas-phase spectra.

Frequencies of *trans*-NMA hydrogen bonded with two water molecules via N−H and C═O groups have been calculated by Krimm et al.[34,73] at the Hartree−Fock level. Surprisingly (considering that our calculation fully takes into account two layers of solvent molecules around NMA instead of only two molecules and considering the difference in the level of calculations), the trends obtained here for the amide band separations are very comparable: We find ∼52−69 cm^{-1} between the amide I and amide II bands, while ∼50−70 cm^{-1} is found in refs 34 and 73, and our ∼250−230 cm^{-1} amide II/amide III gap compares very well with the 277 cm^{-1} gap of refs 34 and 73.

7. Infrared Spectroscopy

Unless stated otherwise, the infrared spectra discussed below are calculated using the dipole−dipole correlation function of the eq 1 of section 3. The alternative way of calculating infrared spectra through the current−current correlation function, eq 6 of section 3, will be discussed in section 7.3 for the particular case of aqueous *trans*-NMA.

7.1. Infrared Spectra of Gas-Phase NMA. We have reported in Figure 4 the infrared spectrum of isolated *trans*-NMA calculated from the quantum molecular dynamics trajectory, again in the 1000−2000 cm^{-1} spectral domain. The active infrared modes are identified by comparing to the vibrational density of states (Table 4). A schematic representation of vibrational assignments is presented at the top of Figure 5. Six infrared active bands are observed in the 1000−2000 cm^{-1} region. They are relatively narrow as expected in the gas phase. The most intense IR band located at 1609 cm^{-1} is related to the amide I mode. In the spectral domain between 1500 and 1300 cm^{-1}, we obtain four active infrared bands of much lower intensity: The amide II mode (1458 cm^{-1}), $\delta(C-H)_C$ (1407 cm^{-1}), and $\delta(C-H)_N$ (1367 cm^{-1}) bending, and the backbone stretch $\nu(C-CH_3)$ coupled to $\delta(C-H)_C$ (1327 cm^{-1}). As previously described in the VDOS analysis, the mirror mode on the N-terminal side, i.e., $\delta(C-H)_N/\nu(N-CH_3)$, should contribute to the modulation of the dipole moment of the molecule, thus giving rise to an active infrared band, but it does not. The more intense active band located at 1189 cm^{-1} is due to amide III. The intensity of this band remains nonetheless much smaller than the one of amide I. All other modes related to "pure" motions of the methyl groups (without any strong contributions from the skeleton atoms) do not show any infrared activity. Surprisingly, the collective backbone stretching deformation mode combined with methyl rocking, located at ∼1000 cm^{-1}, has no infrared activity either. The amplitude ratios obtained in our calculated bands are ∼10 for amide I−amide II, ∼5 for amide I−amide III, and ∼0.5 for amide II−amide III.

The result for the infra spectrum of gas-phase *cis*-NMA is given in Figure 4, and a schematic representation of the vibrational assignments appears at the top of Figure 6. We immediately see that this spectrum is totally different from the *trans*-NMA spectrum: in particular, the number of active bands is reduced from six to four, and the 420 cm^{-1} overall frequency range of the *trans*-NMA spectrum is reduced to ∼350 cm^{-1}. The three main infrared bands have respectively a very high (1606 cm^{-1}), high (1331 cm^{-1}), and medium intensity (1259 cm^{-1}). One additional band of very low intensity is located around 1400 cm^{-1}. The most intense active band is associated with the amide I mode. The

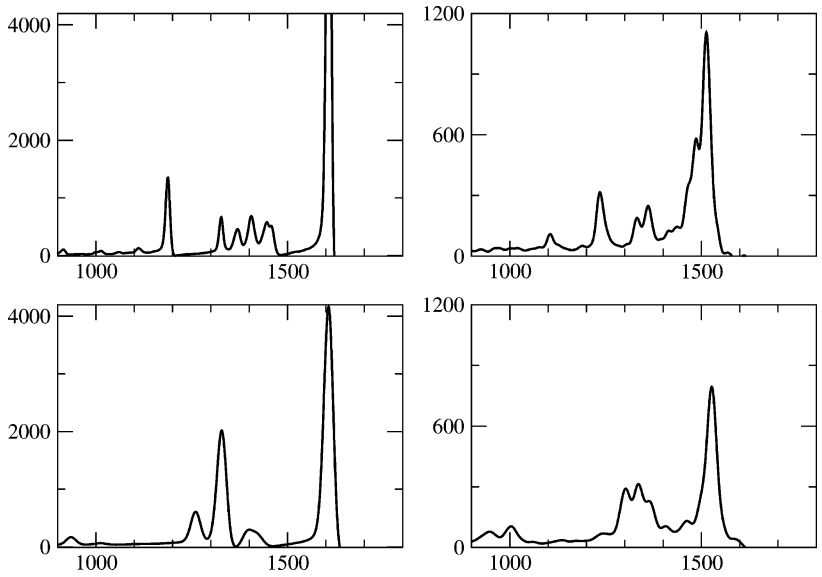

Figure 4. Calculated infrared spectra of gas-phase *trans*-NMA (top left), gas-phase *cis*-NMA (bottom left), aqueous *trans*-NMA (top right), and aqueous *cis*-NMA (bottom right). Spectra are reported in the 1000−1800 cm^{-1} wavenumber range, and IR amplitudes are reported in units of cm^{-1}.

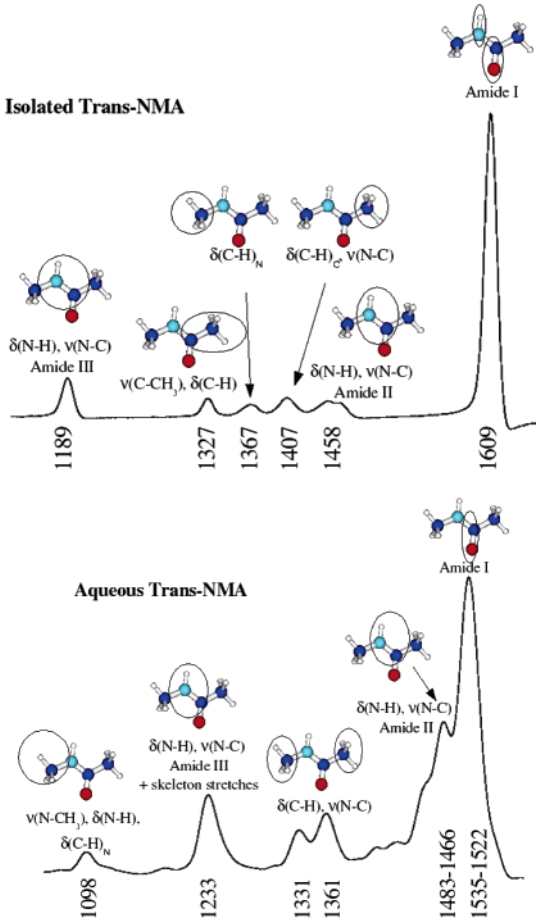

Figure 5. Schematic representation of the assignments of the calculated infrared spectrum of gas-phase (top) and aqueous-phase *trans*-NMA (bottom).

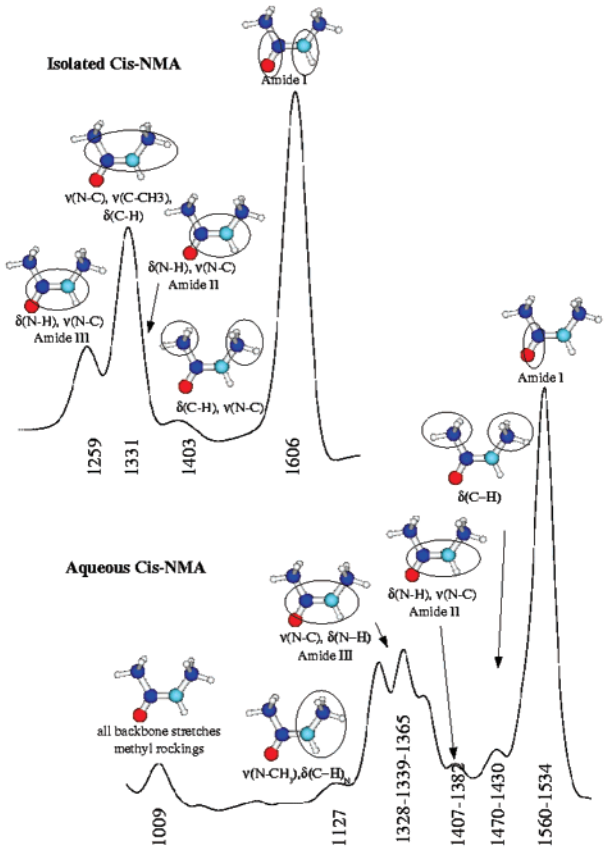

Figure 6. Schematic representation of the assignments of the calculated infrared spectrum of gas-phase (top) and aqueous-phase *cis*-NMA (bottom).

relatively broad band peaked at 1331 cm^{-1} and extending between ∼1300 cm^{-1} and ∼1400 cm^{-1} is composed of the $\nu(C-CH_3)/\delta(C-H)/\nu(N-C)$ mode identified previously at 1331 cm^{-1} in the VDOS, together with the amide II mode (1369 cm^{-1}) in the higher wavenumber part of the band. Thus, in our calculation, amide II does not appear in a single peak but is instead merged into a broader band which combines two vibrational modes. The infrared active mode located at 1259 cm^{-1} is due to amide III. The band of very weak intensity around 1400 cm^{-1} is due to $\delta(C-H)$ coupled

to ν(N−C) motions. Amplitude ratios between the peaks are ∼1.7 between amide I and the 1331 cm^{-1} band and ∼3.3 between amide I and amide III.

The amide I band is thus peaked in the two spectra of *trans*- and *cis*-NMA at the same frequency, and for both molecules, this band has the most intense activity of the whole spectrum. A second remark is that the IR spectrum of *trans*-NMA is composed of four distinct bands of rather low intensity in the 1500−1200 cm^{-1} domain (among which amide II) which are well separated from the more intense amide III band located at 1189 cm^{-1}. In the same domain, we observe only two distinct infrared bands, of respectively intense and medium intensities, in the spectrum of *cis*-NMA. A striking difference is the amide II mode which emerges as a single band in *trans*-NMA, whereas it is merged with the ν(C−CH$_3$) stretching band at 1331 cm^{-1} in *cis*-NMA. We note that the band with maximum at 1331 cm^{-1} in both spectra has a much higher activity in *cis*-NMA, presumably because of the merging with the amide II mode. The two δ(C−H) infrared bands of *trans*-NMA (1407 and 1367 cm^{-1}) are converted into one single peak in *cis*-NMA (∼1400 cm^{-1}) in which the N-terminal and C-terminal methyl groups jointly participate.

7.2. Infrared Spectra of Aqueous NMA. The infrared spectrum as computed for aqueous *trans*-NMA is shown in Figure 4. A schematic representation of vibrational assignments is added at the bottom of Figure 5. In the 1000−2000 cm^{-1} spectral domain the spectrum consists of six main peaks. The band broadly extending between 1450 and 1550 cm^{-1} has the highest intensity in the overall spectrum. Amide I and amide II modes are merged into this band, and amide II appears as a shoulder of high intensity. The infrared peaks of medium intensity located at 1361 and 1331 cm^{-1} come from δ(C−H) vibrations (combined with some ν(N−C) stretch). The peak of high intensity at 1233 cm^{-1} arises from absorption by amide III motion. Its high intensity is the consequence of a strong modulation of the dipole moment of the molecule through the N−C stretching and the combination with other skeleton stretching, evidenced in the VDOS analysis. The intensity of this band is about a third of that of the coupled amide I−amide II one. Finally, the band of low intensity at 1098 cm^{-1} is due to ν(N−CH$_3$)/δ-(C−H)$_N$/δ(N−H).

Turning to aqueous *cis*-NMA the calculated IR spectrum of this conformer in the 1000−2000 cm^{-1} region (Figures 4 and 6) is composed of two main bands positioned at ∼1526 cm^{-1} and around 1339 cm^{-1} and three peaks of lower intensities around 1450, 1400, and 1000 cm^{-1}. We can also distinguish a shoulder of very weak intensity around 1127 cm^{-1}. The very intense band peaking at 1526 cm^{-1} and roughly extending between 1400 and 1600 cm^{-1} is amide I. The intense band centered around 1339 cm^{-1} is resolved in three distinct maxima and is due to a mixing of amide III and ν(C−CH$_3$)/δ(C−H) (modes at 1328−1339−1365 cm^{-1} in the VDOS analysis of Table 7). This band combines on the high wavenumber side with the lower intensity amide II band (1382−1407 cm^{-1} in the VDOS). The shoulder of weak intensity at ∼1127 cm^{-1} is related to ν(N−CH$_3$)/δ(C−H)$_N$, while the collective backbone stretching and methyl rocking give rise to the active band of medium intensity located at ∼1000 cm^{-1}.

Thus, the band features displayed by the infrared spectra of aqueous *trans*- and *cis*-NMA are also quite different, as already observed in the gas phase. In particular, in the IR spectrum of aqueous *trans*-NMA the amide I and amide II modes are combined in a single very intense active band, while two intense bands appear in the spectrum of aqueous *cis*-NMA, i.e., amide I and the merged amide III−amide II. As a consequence, we can see that the 1300−1600 cm^{-1} spectral domain can be used as a very good spectral indicator of the liquid-phase trans versus cis conformation. The 1000−1300 cm^{-1} spectral domain also displays very different features for both conformers. The spectrum of *cis*-NMA appears flat in this region, whereas *trans*-NMA produces two bands, one at 1233 cm^{-1} corresponding to the amide III mode and the other at 1100 cm^{-1} for ν(N−CH$_3$)/δ(C−H)$_N$. On the other hand, the combination of skeleton stretching motions and methyl rocking around ∼1000 cm^{-1} has a substantial amplitude for *cis*-NMA but is almost absent in the spectrum of *trans*-NMA.

Comparing the infrared absorption in the gas and liquid phase, we can infer the following solvent effects on the infrared spectra of *trans*- and *cis*-NMA. Starting with *trans*-NMA, a striking consequence of the interaction with the solvent is the combination of a red-shift of the amide I and blue-shift of amide II band reducing the 150 cm^{-1} frequency gap observed in the gas phase to ∼40−50 cm^{-1} in solution. The result is the broad double band containing both modes. In both gas and liquid phases, apart from the very intense amide I band, the more intense band is due to the amide III mode, and solvation induces a blue-shift of 40 cm^{-1}. A further noticeable solvent effect is the gain of infrared activity of ν(N−CH$_3$)/δ(C−H)$_N$ at 1098 cm^{-1}. For *cis*-NMA, the main changes concern the two most intense bands in the gas-phase infrared spectrum, i.e., amide I at 1606 cm^{-1} and the merging of ν(C−CH$_3$), amide II, and amide III around 1200−1400 cm^{-1}. These bands carry comparable intensities in the gas phase and aqueous spectra. However, amide I is 80 cm^{-1} downshifted by the presence of the solvent, whereas amide III is blue-shifted by the same amount. Since ν(C−CH$_3$) and amide II keep basically the same position, it is the change of amide III which is mostly responsible for the band shape evolution. Note again the gain of intensity of the collective backbone motion at ∼1000 cm^{-1} in solution.

It is also instructive to relate the red- and blue-shifts of the amide bands to the charge flows and resonant states electronic picture described in section 5. The red-shift of amide I mode in the aqueous phase appears as being due to the weakening of the C═O bond which is the result of intermolecular hydrogen bonding and increased weight of the zwitterionic form. On the other hand, blue-shifts of the amide II and amide III bands can be related to the observed strengthening of N−C bond. Polarizations of both N−H and C═O bonds upon hydrogen bonding, which increase the Born charges of all these atoms (in absolute value), seems to play the most important role in the increase of IR intensities of the amide II and amide III bands, particularly with respect to the charges on N and C atoms. These

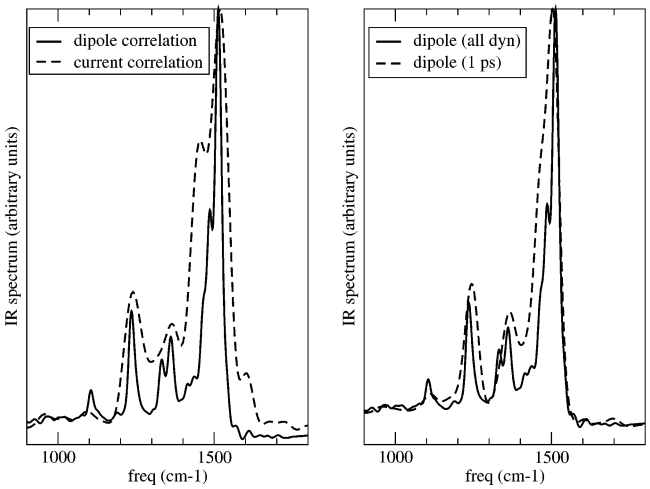

Figure 7. Left: comparison of the NMA infrared absorption spectrum calculated from the solute dipole−dipole correlation function (solid line) obtained over 9.35 ps of dynamics and from current−current correlation function (broken line) obtained over 1.0 ps of dynamics (see section 3). Right: NMA infrared spectrum calculated from the dipole−dipole correlation function over the 9.35 ps of dynamics (solid line) and over 1 ps of dynamics (broken line). See text for conclusions.

intensities however will be very sensitive to the charge transfer from the N-terminal to the C-terminal end, induced by the conjugated character of the N−C═O system, which would tend to lower the charges on N and C and increase (in absolute value) the charges of H and O. A proper account of the electronic structure of the peptide bond is thus needed to accurately reproduce the IR spectra of *cis*- and *trans*-NMA in solution. These observations will also have consequences on the interpretation of the peak shifts of the amide bands of peptides in different environments.

7.3. Dipole−Dipole versus Current−Current Correlation Functions. As explained in section 3, we have developed an alternative scheme for extracting the solute spectrum from the IR absorption of the full solution. This method is based on current−current autocorrelation functions (see eqs 2−6). The current is written in terms of atomic APT tensors calculated through linear response theory and atomic velocities. In Figure 7, we compare the IR spectrum of NMA computed in the dipole representation using Wannier centers to the result of the current representation based on atomic APT tensors. Note that the infrared spectrum determined from Wannier centers has been calculated over the whole ab initio trajectory, whereas only 1 ps of the total trajectory was used when the absorption was computed from current fluctuations. This limitation reflects the extra CPU cost of linear response theory for the calculation of the atomic APT tensors, as already mentioned in section 3. We thus expect that the statistics of the current spectrum will be better for the higher wavenumbers (∼above 1300−1400 cm^{-1}) than in the lower part of the spectrum (mostly below 1300 cm^{-1}). For the convenience of the discussion, the spectra of Figure 7 have been normalized to get the same amplitude of the amide I band. For a complete comparison, we have also shown in the same figure the infrared spectrum of NMA calculated with the dipole−dipole correlation function over a slice of 1 ps of the ab initio trajectory.

Despite the very short time span (1 ps) the IR spectrum determined from the current−current autocorrelation function reproduces most of the important features of the absorption. In particular, all amide bands are present, whereas the most intense amide I−amide II band is clearly not correctly obtained from the same 1 ps time interval when the dipole−dipole correlation function is used. This improved convergence is most likely an effect of the favorable statistics of velocities. Atomic velocities, in contrast to dipoles, are isotropic and fluctuate very quickly during the dynamics. Therefore, calculation of infrared spectra through current−current correlation functions can be done on shorter time scales of dynamics. This might be particularly important in the case of strong coupling between almost degenerate modes, such as for example the δ(O−H) bending mode of water and the amide I and amide II bands of NMA which both occupy the same ∼1600 cm^{-1} frequency band. The good agreement between the dipole and current based infrared spectra also validates our scheme for the computation of atomic APT tensors or, conversely, the method for separation of the total dipole moment in solute and solvent contributions.

8. Comparison to Experiment

Experimental infrared spectra are available only for *trans*-NMA, the most stable isomer in both the gas and liquid phase.

8.1. Infrared Spectra in the Gas Phase. Our calculated infrared spectrum of isolated *trans*-NMA can be compared to the experimental spectra obtained in an argon matrix[27] at 20 K and in the gas phase[7] at ∼100 K. The comparisons are displayed in Figure 8a. Note immediately that the two experimental results are in quantitative agreement, except obviously for substantially broader bands in the gas-phase experiment which was performed at higher temperature. Since the BLYP DFT functional is known to systematically underestimate the experimental frequencies, a conventional practice is to apply a constant multiplicative correction factor to the theoretical results. Here we have taken a factor of 1.064 which adjusts exactly the position of the amide III band to the experimental value of 1266 cm^{-1}. Furthermore the theoretical spectrum has also been normalized as to reproduce the correct intensity of this band. Doing so, we find that the theoretical amide I band is close to the correct position, but its intensity is overestimated. We ruled out that this overestimation be related to a nonthermalized amide I mode. Thermalization of all degrees of freedom can indeed be difficult to achieve within short gas-phase dynamics and can thus induce errors in calculated infrared intensities.[82] The intermediate region between 1300 and 1600 cm^{-1} is only approximately reproduced but with correct features. In light of our previous analysis of the ab initio spectrum, the experimental spectrum in that region can be understood as the succession of the ν(C−CH_3)/δ(C−H)$_C$ band at 1360 cm^{-1}, a series of two (gas phase) to four (matrix) low intensity mixed bands due to CH_3 bends, and the more intense amide II band. In the theoretical spectrum, the ν-(C−CH_3)/δ(C−H)$_C$ band is upshifted and slightly too weak,

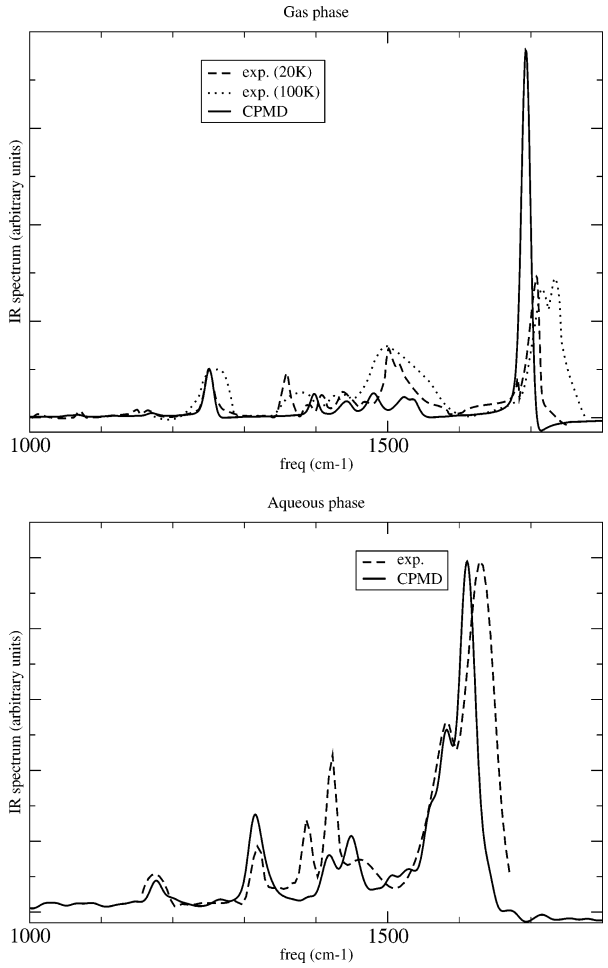

Figure 8. Comparison of the calculated and experimental infrared spectra of *trans*-NMA. Gas phase (top): Car−Parrinello calculation at 20 K (line), experiment in an argon matrix at 20 K[27] (dashed line), experiment in the gas phase at 100 K[7] (dotted line); we note that amplitude of the amide I band in ref 27 has been truncated so that the value is not known. Liquid phase (bottom): Car−Parrinello calculation at 320 K (solid line) and the experiment of refs 32 and 31 at 300 K (dashed line). See text section 8 for the explanation of adjustments made for the sake of comparison between simulation and experiment. See also text for conclusions.

and the amide II band is well positioned but its intensity is definitely too weak. Overall however, bearing in mind the known deficiency of the BLYP functional and the rather simple quantization procedure which was adopted and relies on a picture of coupled harmonic oscillators, it is fair to say that the performance of the ab initio simulations is rather satisfactory, even in the 1300−1600 cm^{-1} spectral range which is quite congested and involves a lot of coupled (anharmonic) motions.

8.2. Infrared Spectra in the Solution. Next we compare the calculated infrared spectrum of aqueous *trans*-NMA to experiments by Song et al.[32,31] and Keiderling et al.[7] All three experiments give very similar results (see Table 5), with subtle differences in band intensities. The experimental spectra presented in refs 31 and 32 have been recorded in the 1000−2000 cm^{-1} spectral domain, while the one of ref 7 is restricted to the 1200−2000 cm^{-1} domain. In Figure 8b, the data of Asher's et al.[31,32] have been used. Again, the computed frequencies have been uniformly scaled by a factor of 1.117 in order to align the amide III band with the experimental peak at ∼1318 cm^{-1}, and we have applied a normalization factor in order to reproduce the intensity for the amide I peak. It can be seen immediately that the overall agreement is very satisfactory with the correct number of bands and correct band shapes, but some discrepancies remain that are worth commenting in the light of our band assignments of Table 5 and Figure 5.

The broad and intense double band in the 1500−1700 cm^{-1} range due to the mixing of amide I and amide II is very well reproduced, except for a slightly lower amide I/amide II gap compared to experiment (54 instead of 60 cm^{-1}). The experimental spectrum exhibits three overlapping bands between 1350 and 1450 cm^{-1} and so does the computed one. In this triplet which was attributed to mostly δ(C−H) bending (coupled to some ν(N−C)), the two first bands appear with the correct relative intensity ratio, but they are slightly too weak and upshifted by ∼40 cm^{-1}. The third one, which in the experimental spectrum contributes to the triplet as a shoulder of lower intensity, is slightly shifted upward to become a shoulder of the amide II peak in the theoretical spectrum. Finally, the computed amide III band is found with a slightly too high intensity, and the lowest frequency band around 1170 cm^{-1}, attributed to ν(N−CH_3)/δ(C−H)$_N$/δ(N−H), is very well reproduced.

We conclude that the comparison of our calculations of the aqueous *trans*-NMA IR absorption to experiments gives overall excellent agreement in terms of the number of active bands, relative positions, and band shapes. The main discrepancy comes from the two twin bands related to δ(C−H) bending, for which one gets a spurious blue-shift and too weak intensities. Note that those two trends were also observed in the theoretical gas-phase spectrum. Comparison of our spectrum to the one simulated by Besley[35] using harmonic frequencies calculated over several thousand optimized structures (ab initio level) of NMA with three explicit water molecules embedded in a dielectric continuum, shows a better agreement to experiment. In particular, we obtain a better agreement on the amplitude of the amide II band, while the methyl bands lack any activity in ref 35. As in our calculated spectrum, the intensity of the amide III band is slightly overestimated in Besley's work.[35] We note that Besley's calculations[35] remain rather more cumbersome (several thousand conformations optimized) and expensive in comparison to our ab initio molecular dynamics. Apart from CH_3 bending, the comparison of our calculated infrared spectrum of aqueous *trans*-NMA with the one calculated by Kato et al.[37] using a classical polarizable force field shows that our quantum molecular dynamics simulations are doing much better. In Kato's work, the general features of the amide I−amide II double bands are well represented, though the amplitude of the amide II band is clearly underestimated. Nonetheless, all other infrared bands in the 1000−1500 cm^{-1} domain below this band are lacking activity. For example, the two peaks related to δ(C−H) bending have no infrared intensity, and the amide III band is barely discernible in the spectrum. Moreover, the weak band showing up in our

calculation and in experiments at $\sim$1098 cm^{-1} is missing in Kato's calculation. Electronic polarization when treated at the ab initio level, as in Car−Parrinello calculations, gives thus a better insight into the infrared patterns. However, a number of problems remain to be resolved, such as underestimation and misplacement of CH_3 infrared bands.

9. Summary and Conclusions

In this paper, we have studied the characteristic vibrational features of *N*-methylacetamide, a simple but essential building block of peptides and proteins, using Car−Parrinello ab initio molecular dynamics simulations. The infrared spectra of *trans*-NMA and *cis*-NMA in the gas and aqueous phase were analyzed and compared in detail. The key computational tool for the calculation of molecular dipoles is the maximally localized Wannier functions. As demonstrated in a previous study of aqueous uracil,[12] and confirmed here, the Wannier functions approach leads to accurate infrared spectra which can be directly compared to experiments. Although band positions are systematically red-shifted (mainly due to our use of the BLYP functional), the main infrared vibrational features, including band shapes and relative bands gaps, can be satisfactorily reproduced. The focus in the present NMA study was on the amide bands. Our results for the relative position and amplitude of the amide I and amide II are in very good agreement with experiments. Amide III also ends up in the right position relative to amide I and II but with a slightly overestimated intensity. It is thus fair to say that our methodology is able to reproduce correctly all amide features in the 1000−2000 cm^{-1}, which is indeed an essential prerequisite to the future study of the vibrational spectroscopy of larger polypeptides and proteins. The amplitudes of methyl groups δ(C−H) bands were underestimated in our calculations. Since C−H/water interactions are more sensitive to dispersion than to electrostatics forces, this deficiency could be related to the lack of a proper dispersion term in DFT calculations.

To compare our calculated infrared spectra to the experiments (gas and liquid phase), we made use of two different scaling factors that adjusted the position of the calculated amide III band (1.064 and 1.117, respectively, for the gas phase and the solution). This point deserves some more comments. There is no real reason as to why the scaling factor between gas phase and solution should be rigorously the same. This scaling factor depends on frequency and thus could change in a condensed phase environment. What could also be different in solution is an enhanced inertia (giving rise to frequency red-shifts) due to the fictitious electron mass used in the Car−Parrinello molecular dynamics scheme. As we have shown in our previous paper,[12] the fictitious electron mass can contribute to the underestimation of the frequencies (up to 40−50 cm^{-1} for the study[12]). Added to the well-known frequency red-shifts due to the use of the BLYP functional, this leads to an increased underestimation of the frequency positions.

An important technical issue that concerned us in this paper was the decomposition of the total spectrum into solute and solvent components. We have used two alternative schemes to achieve such a separation. One way, already mentioned above, is based on the resolution of the cell dipole moment in molecular dipoles computed from maximally localized Wannier functions and the subsequent calculation of the infrared spectrum by Fourier transformation of the solute dipole−dipole correlation function. This strategy is the same as adopted in our previous work.[12] An alternative route, applied here for the first time for a solute in a solvent, involves the representation of the IR signal in terms of a current correlation function and the separation of the current into atomic components. The atomic charges used in the expression of the current were obtained from atomic polar tensors (APT's) calculated using linear response methods. We showed that both definitions yield very similar results. The relative merits of the two methods deserve some comment. A clear drawback of the current based scheme is that the computation of APT tensors is expensive. However, the definition of molecular dipole moments is not unique. The definition of molecular currents is relatively straightforward and not subject to the same arbitrariness as the dipoles. Moreover, such an analysis seems to yield meaningful results with relatively short dynamics, at least shorter than those required with the Wannier decomposition.

Our results show how the activity of the infrared bands of the two NMA isomers is affected by the liquid phase. We were also able to relate these shifts of the amide bands to the solute−solvent intermolecular hydrogen bond network. The study of the Wannier centers and, in the case of *trans*-NMA, of the Born charges were found to provide a useful physical picture of the charge rearrangement in response to solvation. The electrons flows observed in the HNCO peptide link are consistent with an increased weight of the zwitterionic form in a two-state resonance model. This relates the solvent-induced spectral changes to the weakening of the C═O bond and strengthening of the N−C bond. We believe that these observations will be useful for further modeling of the IR spectra (amide bands) of larger peptides in different environment and/or conformations using molecular force fields.

Our simulations also emphasize the importance of the amide III band among the infrared signatures. In particular, the blue-shift of this band in going from the gas to the liquid phase can lead to specific spectral patterns in the IR spectra, as observed for aqueous *cis*-NMA. Interpretation of this band is the least understood of amide bands of peptides and proteins[74−76] and is thought to be highly dependent on the Φ/Ψ angles. We hope that further Car−Parrinello dynamics on more realistic models of peptides will contribute to give a better understanding.

Finally, by comparing the infrared spectra of *trans*-NMA and *cis*-NMA in water determined at the ab initio level (in the DFT sense), we predicted that the two isomers can be clearly distinguished by their IR signatures. In particular the amide I and amide II bands display characteristic differences, both in the gas and aqueous phase. We hope that our findings for NMA can be confirmed experimentally. We also hope that a similar approach can be extended in the future to various helix or β-sheetlike conformations of polypeptides, with the ultimate goal of contributing to the understanding

of how vibrational spectroscopy can help to discriminate protein structures.

Acknowledgment. The authors thank CINES (Montpellier, France) and IDRIS (Orsay, France) for generous access to their computational facilities. The linear response calculations were performed using the HPCx facilities at Daresbury Laboratories as part of the UKCP grant. M.P.G. acknowledges support from Genopole-France through the action 'ATIGE' Action Thématique Incitative de Génopole.

Supporting Information Available: Analysis of the vibrational density of states (VDOS) of *trans*- and *cis*-NMA inferred from the Car−Parrinello trajectories of the isomers in both the gas- and liquid-phase is summarized in Figures 1−4. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Krimm, S.; Bandekar, J. *Adv. Prot. Chem.* **1986**, *38*, 181.

(2) Torii, H.; Tasumi, M. *J. Chem. Phys.* **1992**, *96*, 3379.

(3) *Biological Applications of Raman Spectroscopy;* Spiro, T., Ed.; Wiley-Interscience: New York, 1987; Vol. I.

(4) *Infrared Spectroscopy of Biomolecules*; Mantsch, H., Chapman, D., Eds.; Wiley-Liss: New York, 1996.

(5) Kato, K.; Matsui, T.; Tanaka, S. *Appl. Spectrosc.* **1987**, *41*, 861.

(6) Pancoska, P.; Kubelka, J.; Keiderling, T. A. *Appl. Spectrosc.* **1999**, *53*, 655.

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

(8) Gnanakaran, S.; Hochstrasser, R. M. *J. Am. Chem. Soc.* **2001**, *123*, 12886.

(9) Hamm, P.; Lim, M.; Hochstrasser, R. M. *J. Phys. Chem. B* **1998**, *102*, 6123.

(10) Woutersen, S.; Hamm, P. *J. Phys. Chem. B* **2000**, *104*, 11316.

(11) Moran, A. M.; Dreyer, J.; Mukamel, S. *J. Chem. Phys.* **2003**, *118*, 1347.

(12) Gaigeot, M.-P.; Sprik, M. *J. Phys. Chem. B* **2003**, *107*, 10344.

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

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

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

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

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

(18) Polavarapu, P. L.; Deng, Z.; Ewig, C. S. *J. Phys. Chem.* **1994**, *98*, 9919.

(19) Mirkin, N. G.; Krimm, S. *J. Am. Chem. Soc.* **1991**, *113*, 9742.

(20) Mirkin, N. G.; Krimm, S. *J. Mol. Struct. (THEOCHEM)* **1991**, *236*, 97.

(21) Cuevas, G.; Renugopalakrishnan, V.; Madrid, G.; Hagler, A. T. *Phys. Chem. Chem. Phys.* **2002**, *4*, 1490.

(22) Gregurik, S. K.; Chaban, G. M.; Gerber, R. B. *J. Phys. Chem. A* **2002**, *106*, 8696.

(23) Gao, J.; Freindorf, M. *J. Phys. Chem. A* **1997**, *101*, 3182.

(24) Garcia-Martinez, A.; Teso-Vilar, E.; Garcia-Fraile, A.; Martinez-Ruiz, P. *J. Phys. Chem. A* **2002**, *106*, 4942.

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

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

(27) Ataka, S.; Takeuchi, H.; Tasumi, M. *J. Mol. Struct.* **1984**, *113*, 147.

(28) Fillaux, F.; De Loze, C. *J. Chim. Phys.* **1976**, *73*, 1004.

(29) Grenie, Y.; Avignon, M.; Garrigou-Lagrange, C. *J. Mol. Struct.* **1975**, *24*, 293.

(30) Chen, X. G.; Schweitzer-Stenner, R.; Krimm, S.; Mirkin, N. G.; Asher, S. A. *J. Am. Chem. Soc.* **1994**, *116*, 11141.

(31) Song, S. S.; Asher, S. A.; Krimm, S.; Bandekar, J. *J. Am. Chem. Soc.* **1988**, *110*, 8547.

(32) Song, S. S.; Asher, S. A.; Krimm, S.; Shaw, K. D. *J. Am. Chem. Soc.* **1991**, *113*, 1155.

(33) Herrebout, W. A.; Clou, K.; Desseyn, H. O. *J. Phys. Chem. A* **2001**, *105*, 4865.

(34) Chen, X. G.; Schweitzer-Stenner, R.; Asher, S. A.; Mirkin, N. G.; Krimm, S. *J. Phys. Chem.* **1995**, *99*, 3074.

(35) Besley, N. A. *J. Phys. Chem. A* **2004**, *108*, 10794.

(36) MacKerrel, A. D.; Karplus, M. *J. Am. Chem. Soc.* **1995**, *117*, 11946.

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

(38) Mantz, Y. A.; Gerard, H.; Iftimie, R.; Martyna, G. J. *J. Am. Chem. Soc.* **2004**, *126*, 4080.

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

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

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

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

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

(44) Hutter, J. et al. *CPMD, version 3.7.1*. IBM Research Division, IBM Corp and Max Planck Institute Stuttgart.

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

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

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

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

(49) Cioslowski, J. *J. Am. Chem. Soc.* **1989**, *111*, 8333.

(50) Gonze, X. *Phys. Rev. B* **1997**, *55*, 10337.

(51) Wilson, E. B.; Decius, J. C.; Cross, P. C. *Molecular Vibrations*; McGraw-Hill: New York, 1955.

(52) De Proft, F.; Martin, J. M. L.; Geerlings, P. *Chem. Phys. Lett.* **1996**, *250*, 393.

(53) Pasquarello, A.; Car, R. *Phys. Rev. Lett.* **1997**, *79*, 1766.

(54) Pasquarello, A.; Resta, R. *Phys. Rev. B* **2003**, *68*, 174302.

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

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

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

(58) Blondel, A.; Karplus, M. J. *J. Comput. Chem.* **1996**, *17*, 1132.

(59) Kitano, M.; Fukuyama, T.; Kuchitsu, K. *Bull. Chem. Soc. Jpn.* **1973**, *46*, 384.

(60) Du, Q.; Wei, D. *J. Phys. Chem. B* **2003**, *107*, 13463.

(61) Gaigeot, M.-P.; Sprik, M. *J. Phys. Chem. B* **2004**, *108*, 7458.

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

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

(64) Izvekov, I.; Voth, G. A. *J. Chem. Phys.* **2002**, *116*, 10372.

(65) Grossman, J. C.; Schwegler, E.; Draeger, E. W.; Gygi, F.; Galli, G. *J. Chem. Phys.* **2004**, *120*, 300.

(66) VandeVondele, J.; Mohamed, F.; Krack, M.; Hutter, J.; Sprik, M.; Parrinello, M. *J. Chem. Phys.* **2005**, *122*, 014515.

(67) Torii, H.; Tatsumi, T.; Tasumi, M. *J. Raman Spec.* **1998**, *29*, 537.

(68) Rodrigo, M. M.; Tarazona, M. P.; Saiz, E. *J. Phys. Chem.* **1986**, *90*, 2236.

(69) Pralat, K.; Jadzyn, J.; Balanicka, S. *J. Phys. Chem.* **1983**, *87*, 1385.

(70) Wang, Y.; Purrello, R.; Jordan, T.; Spiro, T. G. *J. Am. Chem. Soc.* **1991**, *113*, 6359.

(71) Mayne, L. C.; Hudson, B. *J. Phys. Chem.* **1991**, *95*, 2962.

(72) Miyazawa, T.; Shimanouchi, T.; Mizushima, S. I. *J. Chem. Phys.* **1956**, *24*, 408.

(73) Mirkin, N. G.; Krimm, S. *J. Mol. Struct.* **1996**, *377*, 219.

(74) Mikhonin, A. V.; Ahmed, Z.; Ianoul, A.; Asher, S. A. *J. Phys. Chem. B* **2004**, *108*, 19020.

(75) Mirkin, N. G.; Krimm, S. *J. Phys. Chem. A* **2002**, *106*, 3391.

(76) Schweitzer-Stenner, R.; Eker, F.; Huang, Q.; Griebenow, K.; Mroz, P. A.; Kozlowski, P. M. *J. Phys. Chem. B* **2002**, *106*, 4294.

(77) Ramirez, R.; Lopez-Ciudad, T.; Kumar, P.; Marx, D. *J. Chem. Phys.* **2004**, *121*, 3973.

(78) Person, W. B.; Newton, J. H. *J. Chem. Phys.* **1974**, *61*, 1040.

(79) Person, W. B. In *Vibrational Intensities In Infrared and Raman Spectroscopy;* Person, W. B., Zerbi, G., Eds.; Elsevier: 1982.

(80) Kubelka, J.; Keiderling, T. A. *J. Am. Chem. Soc.* **2001**, *123*, 6142.

(81) Kuo, I.-F. W.; Mundy, C.; McGrath, M. J.; Siepmann, J. I.; VandeVondele, J.; Sprik, M.; Hutter, J.; Chen, B.; Klein, M. L.; Mohamed, F.; Krack, M.; Parrinello, M. *J. Phys. Chem. B* **2004**, *108*, 12990.

(82) Schmitz, M.; Tavan, P. *J. Chem. Phys.* **2004**, *121*, 12233.

CT050029Z