Alanine Polypeptide Structural Fingerprints at Room Temperature: What Can Be Gained from Non-Harmonic Car−Parrinello Molecular Dynamics Simulations

M.-P. Gaigeot

LAMBE UMR8587 Laboratoire Analyse et Modélisation pour la Biologie et l'Environnement, Université d'Evry val d'Essonne, Boulevard F. Mitterrand, Bâtiment Maupertuis, 91025 Evry, France

Received: August 23, 2008; Revised Manuscript Received: October 6, 2008

Structural infrared fingerprints of neutral gas phase alanine peptides of increasing size and complexity (dipeptide, octapeptide, and β-strand peptide) are characterized through DFT-based Car−Parrinello molecular dynamics simulations. Harmonic and nonharmonic vibrational signatures are calculated from the time correlation of the dipole moment of the gas phase peptide in a direct way (without any approximation) respectively from low temperature (20 K) and room temperature (300 K) molecular dynamics. Our main purpose is to answer the two following questions: (i) Is the direct inclusion of temperature for the calculation of infrared spectra mandatory for the comprehension of the vibrational signatures experimentally recorded at room temperature? (ii) To what extent is the amide I, II, and III domain sensitive enough to the local structure of the peptides, to provide vibrational signatures that can be definitely used to assess the peptide conformation at 300 K?

1. Introduction

Vibrational spectroscopy is well established as a probe of polypeptide and protein secondary structures.[1−5] This technique has been largely applied in the liquid phase,[6−13] and the past decade has witnessed the development of several experimental setups in order to probe the vibrational properties of peptides in the gas phase.[14] These experiments can mainly be separated into two groups, with low temperature gas phase infrared spectroscopy on the one hand[14−18] and room temperature gas phase spectroscopy (or close enough to it) based on "action spectroscopy" on the other hand, either using a messenger method (IR-PD, infrared photon dissociation)[19−22] or multiphoton dissociation (IR-MPD, infrared multiphoton dissociation).[23−28] These experiments have been applied to the determination of the structure of biomimetic molecules and peptides.[25−27,29−35] The obvious advantage of gas phase spectroscopy over liquid phase spectroscopy is the direct probe of the intrinsic properties of the molecules, without interference and complications arising from interactions with the surrounding solvent environment. Studies of unsolvated peptides and proteins are of interest not just for fundamental reasons. While the interactions with the solvent are obviously important, as water is an ubiquitous environment of biomolecules, the solvent is nonetheless excluded in many environments such as in the hydrophobic interior of folded proteins or is highly shielded as in membranes.

Vibrational measurements are generally complemented with theoretical calculations in order to get a precise understanding of the vibrational features at the microscopic level. To that end, quantum chemistry calculations are usually performed where geometry optimizations and subsequent vibrational analyses in the harmonic approximation are achieved. Inclusion of anharmonic corrections can be done in principle, but is scarcely applied.[36,37] A match between experimental and calculated patterns obtained for the different isomers is searched, in terms of positions and relative intensities of the spectral bands, in order to provide answers as to which isomer(s) can be responsible for the experimental vibrational signatures. There are several issues in this scheme, though: (i) the determination of the lowest energy equilibrium conformations of the peptide, which can be a tremendous amount of work for floppy peptides; (ii) the lack of vibrational anharmonicities; (iii) temperature, which is not included in static calculations. Temperature is introduced only a posteriori when considering the Boltzmann population factor taking into account the relative weight of the different isomers in the final spectrum. However, temperature plays an important role in the conformational dynamics of floppy molecules, and taking this into account could be pivotal for a better understanding of vibrational patterns. Obviously, the explicit introduction of temperature and the related conformational dynamics for the vibrational signatures is certainly not so useful as long as the calculations are applied for the interpretation of vibrational experiments conducted at low temperature. One can then indeed reasonably admit that the peptide is frozen in structures of lowest energy, and that there is not enough energy (or temperature) to overcome the energy barriers between conformers. This assumption is not valid anymore when the experiments are conducted at room temperature, as is the case for IR-PD and IR-MPD spectroscopies. Room temperature should be enough to overcome the low barriers between isomers of floppy polypeptides, thus resulting in a conformational dynamics which will in turn be pivotal for the vibrational patterns, as we have shown in the case of the protonated alanine dipeptide Ala-Ala-H^+.[35,38]

The finite temperature present in the IR-PD and IR-MPD experiments makes it thus necessary to apply theoretical tools which take temperature into account, in order to accurately interpret the vibrational signatures. Molecular dynamics (MD) simulations at finite temperature are the proper tool to investigate all these properties.[39,40] Taking into account the dynamics of the molecules and its consequences on the measured properties, such as vibrational spectra, can indeed only be achieved through MD simulations. The calculation of infrared spectra through MD relies on dipole time correlation functions recorded along the trajectory.[41] This is well established in the context of classical MD,[42−47] but is much more recent in the ab initio MD community. Such advances were made possible, for instance,

* E-mail: mgaigeot@univ-evry.fr.

10.1021/jp807550j CCC: $40.75
Published on Web 11/19/2008

by the latest developments in the modern theory of polarization and the maximally localized Wannier functions,[48] as implemented within the Car−Parrinello framework.[49−52] Pioneering studies were done by Parrinello et al. on liquid water.[50−52] Within the past few years, we have shown that ab initio molecular dynamics is the proper method for the calculation of IR spectra of DNA and peptide building blocks, in the gas phase or immersed in liquid water,[35,38,53−55] at room temperature. Other authors have more recently applied the same formalism.[56] In particular, our investigation on the gas phase protonated alanine dipeptide Ala-Ala-H^+ [38] demonstrated that CPMD (Car−Parrinello molecular dynamics) simulations are the proper tool to calculate IR absorption spectra of gas phase molecules undergoing multiple isomeric conformations at finite temperature. The reason for that is that all conformations populated when hopping between basins on the potential energy surface during dynamics are naturally accounted for in the calculation of the infrared spectrum in MD.

Another advantage of MD simulations for the calculation of IR spectra is that vibrational anharmonicities are taken into account in the final spectrum, in a direct way. The two successive harmonic approximations usually adopted for the determination of IR spectra from static ab initio calculations, i.e., harmonic approximation of the potential energy surface at the optimized geometries and mechanical harmonic approximation for the transition dipole moments, are indeed released in MD, simply because they are not needed. Hence, the finite temperature dynamics takes place on all accessible parts of the potential energy surface, be they harmonic or anharmonic. As the calculation of IR spectra with MD is related only to the time-dependent dipole moment, it thus does not require any harmonic expansion of the transition dipole moments. Therefore, if the dipole moments and their fluctuations are accurately calculated along the trajectory, the resulting IR spectrum should be reliable too, as demonstrated in our previous works.[35,38,53−55]

Finding the structure of polypeptides in the gas phase is currently a very active area,[15−17,25−27,34,35,57−60] where there is a particular interest in bringing into light helix and sheet secondary structures. In the absence of a solvent, charge plays a critical role in stabilizing helices and destabilizing sheets, as shown by Jarrold's group with ion mobility experiments:[57] β-sheet structures are stable for neutral peptides in vacuo, but are disrupted by a charge; on the contrary, charge stabilizes helical conformations, and helices have not been observed for neutral peptides. Demonstration of these properties through in vacuo infrared spectroscopy is actively being sought, and in this context the amide I−III (1000−2000 cm^{-1}) domain is considered most relevant as a structural probe. From the theoretical point of view, a number of groups have reported harmonic frequency calculations on polypeptides of various lengths and complexity. Krimm and colleagues[1] were the precursors, performing full ab initio geometry optimizations and harmonic frequencies on short peptide models, and applying the TDC model (transition dipole coupling) on long peptide chains. The latest theoretical studies on the harmonic vibrations of long chain polypeptides mainly make use of the DFT (density functional theory) representation,[10,12,61−64] with a most solely emphasis on the amide I band (C═O stretchings). These calculations rely on parameter transfers from small peptides to larger ones. Besley et al.[61] have developed a partial Hessian matrix approximation which allows calculation of harmonic amide I bands of long polypeptides, accurately and with a substantial reduction in CPU time.

The present study is dedicated to the characterization of the vibrational signatures of neutral alanine peptides of increasing size and complexity with ab initio molecular dynamics simulations, with a special emphasis on the effect of temperature and anharmonicities on the vibrational fingerprints. This is a theoretical investigation where we seek to characterize the vibrational signatures of peptides of different lengths and structural complexities, and how the vibrational signatures of the different peptides are modified by temperature. We work on the alanine dipeptide Ala_2, the octaalanine peptide Ala_8, and an alanine β-sheet model (build on two alanine dipeptides held together by intermolecular hydrogen bonds). All peptides are capped by methyl groups at their extremeties. Our goal is to unravel the relations between the amide bands and the peptide structures and dynamics, and demonstrate the pivotal role of temperature and anharmonicities in the IR signatures. Our calculations are also directed at the evolution of the IR signatures with the increase of geometric complexity of the alanine polypeptide. To that end, we have performed ab initio MD simulations at low temperature (20 K, which gives rise to harmonic dynamics and harmonic vibrational signatures which are equivalent to the ones obtained in standard static calculations) and at room temperature (300 K, where anharmonicities are now included in the dynamics and in the vibrational signatures). The 300 K temperature dynamics give us the opportunity to explore the conformational dynamics of the polypeptides and the direct consequences on the infrared spectra. We investigate the three amide bands (amide I, amide II, and amide III) in the 1000−2000 cm^{-1} domain. Our main purpose is to answer the two following questions: (i) Is the direct inclusion of temperature for the calculation of IR spectra mandatory for the comprehension of the vibrational signatures? (ii) To what extent is the amide I, II, and III domain sensitive enough to the local structure of the peptides, to provide vibrational signatures that can be definitely used to assess the peptide conformation at 300 K?

The paper is organized as follows. Section 2 is dedicated to the technical presentation of the DFT-based Car−Parrinello molecular dynamics simulations performed here. Section 3 presents the dynamics and infrared spectra of the polypeptides at low and room temperatures. Conclusions and discussions are presented in section 4. Throughout the text, we will use "red shift" as a means for indicating "a shift toward lower frequency" and "blue shift" as a means for indicating "a shift toward higher frequency".

2. DFT-Based Car−Parrinello Molecular Dynamics

We perform Car−Parrinello molecular dynamics simulations (CPMD),[49] where nuclei are treated classically and electrons quantum mechanically within the DFT (density functional theory) formalism. CPMD simulations consist in solving the nuclei Newton's equations of motions at finite temperature, while the electronic wave function is simultaneously adiabatically propagated. Forces that act on the nuclei derive from the Kohn−Sham energy. CPMD combines plane-wave/pseudopotential methods for the determination of the electronic structure of extended systems within the DFT generalized gradient approximation (GGA) developed for treating molecules. A special feature distinguishing the CPMD method from other ab initio MD methods is the dynamical scheme adopted for the optimization of the electronic states which is based on a fictitious (classical) dynamics of the electronic degrees of freedom.[49,65]

Conditions of the Car−Parrinello simulations performed in this work follow the general setup of our previous ab initio molecular dynamics simulations.[38,53,54,66] 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[67] 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). Hydrogen atoms are represented by a simple s potential damping the Coulomb singularity at the origin. Hydrogen atoms are treated as classical particles with their true mass (1836 au). Energy expectations are calculated in reciprocal space using the Kleinman−Bylander transformation.[68] We used the Becke, Lee, Yang, and Parr (BLYP) gradient-corrected functional[69,70] for the exchange and correlation terms. All simulations were carried out with the CPMD ab initio molecular dynamics package.[71]

Molecular dynamics simulations were performed in the NVE microcanonical ensemble, using a fictitious electron mass of 500 au and a time step of 5 au (0.12 fs). We used cubic boxes of 15 Å length in the simulations of the dipeptides and the β-strand model, and 25 Å for the simulation of the octaalanine. The box length values result from single point wave function optimizations where the convergence of the energy has been checked with respect to the box length (see refs 38 and 72 for more details). A convergence criterion of 10^{-4} au on the energy was applied to that end. Gas phase simulations are carried out with the decoupling technique of Martyna and Tuckerman[73] in order to eliminate the effect of the periodic images of the charge density. All dynamics consist in a thermalization period (∼1.0 ps) where a control of temperature has been applied (through velocity rescaling) and a dynamics without any such control (strictly microcanonical dynamics) over which the data have been collected.

We have performed dynamics at 20 K (harmonic dynamics) and at 300 K (fully anharmonic) for all the systems investigated here. At 20 K, we have performed three dynamics of the alanine dipeptide Ala_2, each one beginning from one of the known optimized structures of least energy (i.e., C7eq, C7ax, C5[74−78]), one dynamics of the β-strand model, and one dynamics of Ala_8 initiated with an α-helix conformation. Simulation durations are respectively 7.6 ps for C7eq, 12.5 ps for C7ax, 9.0 ps for C5, 12.0 ps for the β-strand, and 6.0 ps for Ala_8. The average ionic temperatures obtained in the dynamics are 22 K (C7eq), 21 K (C7ax), 23 K (C5), 22 K (β), and 22 K (helix). The conformations obtained at the end of the 20 K dynamics were subsequently heated to 300 K and thermalized at this temperature. We have performed three dynamics of Ala_2, one of the β-strand, and one of the octaalanine. The durations are respectively 15.7 ps for all trajectories of Ala_2, 11.9 ps for the β-strand trajectory, and 13 ps for the octaalanine Ala_8. Average temperatures over these dynamics are 283 K for Ala_2, 313 K for the β-model, and 288 K for Ala_8.

Calculation of the IR absorption coefficient $\alpha(\omega)$ by means of MD makes use of the relation derived from linear response theory involving the Fourier transform of the dipole time correlation function,[41] as described in our previous works:[35,38,53−55]

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

where $\beta = 1/kT$, $n(\omega)$ is the refractive index, c is the speed of light in a vacuum, and V is the volume. $\mathbf{M}$ is the total dipole moment of the system, which is the sum of the ionic and electronic contributions.[51] It is calculated using the modern theory of polarization framework,[79,80] developed by Silvestrelli and Parrinello[50−52] in the context of Car−Parrinello molecular dynamics simulations. 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.[53,54,81] For a complete discussion on quantum corrections, we refer the reader to refs 82 and 83. The IR spectra are given as products $\alpha(\omega)$ $n(\omega)$ expressed in cm^{-1} as a function of reciprocal wavenumber, ω, in cm^{-1}. The spectra have been smoothed with a window filtering applied in the time domain (Gaussian function with 20 cm^{-1} width), in order to remove the numerical noise arising from the finite length of the Fourier transform.

An accurate calculation of infrared spectra is one goal to achieve, the assignment of the active bands into individual atomic displacements or vibrational modes is another one, and this issue is essential to the understanding of the underlying molecular structural and dynamical properties. In molecular dynamics simulations, interpretation of the infrared active bands into individual atomic displacements is traditionally done using the vibrational density of states (VDOS) formalism. The VDOS is obtained by Fourier transformation of the atomic velocity autocorrelation functions:

$$\mathrm{VDOS}(\omega) = \sum_{i=1,N}\int_{-\infty}^{\infty} \langle \mathbf{v}_i(t)\cdot\mathbf{v}_i(0)\rangle\, \exp(\mathrm{i}\omega t)\, \mathrm{d}t \quad (2)$$

where i runs over all atoms of the investigated system. There are no approximations in this formula. In addition, the VDOS can be decomposed according to each atom type in order to get an interpretation of the vibrational bands in terms of individual atomic motions. This is done by restraining the sum over i in eq 2 to the atoms of interest only. The advantage of the VDOS formalism is that all anharmonicities are taken into account. These schemes have been applied in the present investigations.

In our previous applications of Car−Parrinello dynamics to IR spectroscopy (gas phase and liquid phase calculations), we have systematically found that our calculated infrared spectra must be blue-shifted by 100 cm^{-1} so that the whole calculated bands are aligned with their experimental counterparts. This is a result we have obtained in all the 300 K CPMD simulations we have performed up to now, from which the 800−2000 cm^{-1} spectral region has systematically been extracted. Though our CPMD calculations do not give the proper absolute values of band positions, they do nonetheless give the proper account of spacings between the different active bands. We stress again that a global translation is applied to the whole spectrum and not a frequency-dependent scaling factor. This empirical finding is in contrast to static ab initio calculations where a scaling factor is used to correct the theoretical predictions with respect to the observed frequencies. Note that this empirical scaling is used to correct for the harmonic approximation together with the ab initio representation (ab initio level of theory and basis set). The origin of this wavenumber translation in CPMD is at the moment unclear to us. Effects of the fictitious mass, which lead to instantaneous Car−Parrinello forces being different from Born−Oppenheimer ones whatever the fictitious mass value,[84,85] are certainly important, and indeed the blue shift of 100 cm^{-1} can be reduced when performing Car−Parrinello dynamics with smaller fictitious masses for the propagation of the electronic wave function. This is however at the cost of more expensive simulations. However, it has been demonstrated in refs 84 and 85 that the Car−Parrinello forces can be brought into good agreement with the Born−Oppenheimer forces by simply rescaling the ionic masses (leading to dressed atoms). Were it so, this would amount once more to a scaling factor and not a global translation, as empirically obtained in all our previous

calculations. Such rescaling of the atomic masses should then be dependent on the vibrational mode considered, in a way not fully understood at the moment. All spectra presented thereafter have been translated by 100 cm^{-1} to be consistent with our previous findings.

The quality of the potential energy surface is entirely contained in the "ab initio" force field, calculated at the DFT/BLYP level in the present work, as in our previous investigations.[38,53,54] The very good reproduction of the relative positions (and intensities, when they are directly comparable to the experiment[53,54]) of the different active bands in our previous works is a demonstration that this level of theory is enough. The literature is nonetheless documented with reports that hybrid DFT methods such as B3LYP perform very well at spectroscopic predictions, usually better than gradient-corrected functionals like BLYP, sometimes even out-performing more accurate MP2 calculations. However, our experience in the calculation of IR spectra through finite temperature MD simulations using the BLYP functional proves otherwise, at least on weakly interacting floppy peptide building blocks (gas phase or immersed in liquid water). In ref 38, we have calculated the IR spectrum of the protonated alanine dipeptide with static calculations and with MD simulations, at the same DFT/BLYP level of theory. Our results show that the DFT/BLYP level of theory is not the reason why the static calculations struggle to give an interpretation of the experimental spectrum, and the reason for the spectrum extracted from the DFT/BLYP dynamics to agree with the experiment is entirely contained in the conformational dynamics that is explored during the dynamics. Note finally that the B3LYP hybrid functional has been recently implemented in the CPMD code[86] and in the CP2K MD package,[87] but the CPU cost is reported to be ~40−100 times more expensive than a local functional MD, on massively large parallel supercomputers. Considering our previous publications on the subject together with the extra-CPU cost for a hybrid DFT-based MD, we chose to keep to local BLYP MD simulations in the present work.

3. Results

3.1. Dynamics and Harmonic IR Spectra at 20 K. Our 20 K molecular dynamics have been performed on the alanine dipeptide Ala_2 in its folded C7eq and C7ax and unfolded C5 conformations, which are known to be the lowest energy conformations of Ala_2 (see, for instance, refs 74−78), on a helix octaalanine peptide Ala_8, and on an antiparallel β-strand $Ala_2\cdots Ala_2$ model. All models are capped with methyl groups at the extremeties; see Figure 1. Structures of the dipeptides and the β-strand $Ala_2\cdots Ala_2$ remain in their original symmetry along the 20 K trajectories. We find that the octaalanine peptide adopts a 3_{10} helix at 20 K: the dynamics began with the $i \rightarrow i + 4$ α-helix, and evolved immediately toward the $i \rightarrow i + 3$ 3_{10} helix, which is therefore found here energetically more stable. The notation signifies that the helix is formed through the H-bonded peptide groups i and $i + 4$ ($i + 3$) along the chain. This is in agreement with the results of Keiderling et al.[12] with DFT geometry optimizations of the alanine decamer. See Table 1 reporting the mean values of the backbone Φ and Ψ angles along the trajectories.

On average (see bottom of Table 1), the O···H H-bond distances between the C-terminal C═O and the N-terminal N−H groups along the dynamics are 2.08 ± 0.05 and 1.93 ± 0.05 Å, respectively, for C7eq and C7ax conformations. The shorter H-bond length in C7ax suggests a slightly stronger hydrogen bonding. Note that the average N−H···O H-bond angle of

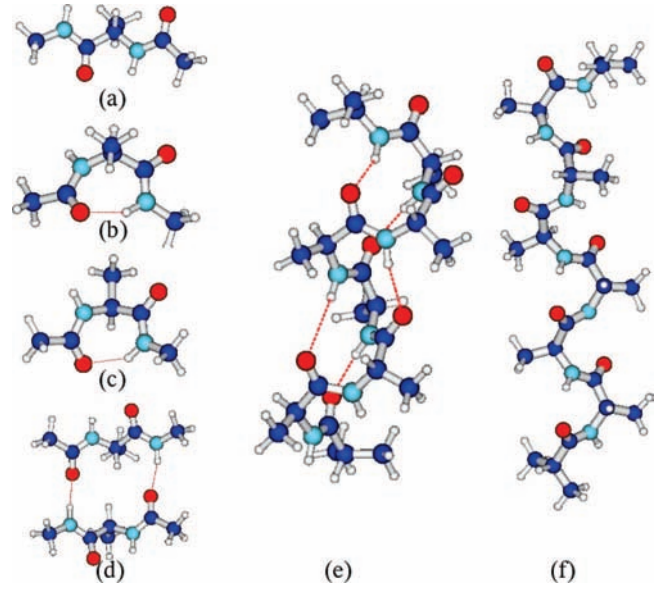

Figure 1. Illustrations of the geometries of the unfolded C5 (a), folded C7ax (b), and folded C7eq (c) of gas phase Ala_2 alanine dipeptide, β-strand alanine model (d), and helix gas phase Ala_8 octaalanine peptide model (e, f). Two structures are represented for Ala_8: the 20 K 3_{10} helix and the average elongated structure extracted from the 300 K dynamics (see text for explanations).

~145−150° obtained for both C7eq and C7ax is the signature of a distorted H-bond. The N-terminal C═O and the C-terminal N−H groups of the C7 conformers do not interact as shown by the average O···H distances along the dynamics, i.e., 3.67 ± 0.08 and 3.99 ± 0.07 Å, respectively, for C7eq and C7ax conformations. The antiparallel β-strand peptide model is maintained over the length of the simulation by two interstrand hydrogen bonds involving the C-terminal C═O and the N-terminal N−H groups of each strand. The average O···H distances of both H-bonds are equal (1.97 Å), which corresponds to energetically equivalent H-bonds. The Φ and Ψ angles of each strand of the β model are very close to the values displayed by the C5 conformer, which confirms that the β peptide can be modeled by C5 conformers linked through interstrand H-bonds. We find five hydrogen bonds between carbonyl and amide groups for the octaalanine 3_{10} helix. The O···H distances range from 2.39 to 2.71 Å (with 0.3−0.5 Å fluctuations), which again corresponds to long and distorted H-bonds. Note that the two carbonyls at the N-terminal of the chain and two amides at the C-terminal of the chain are never involved in H-bonding during the dynamics. Our O···H values are systematically 0.1−0.2 Å higher than the ones obtained by Keiderling et al.[12] with another level of theory.

The average dipole moments of each conformer investigated here at 20 K are reported in Table 2. As expected for low temperature dynamics, the standard deviations from the means are very low. The folded C7ax conformer displays the highest dipole among the two folded conformations, which could be the result of the stronger intramolecular H-bond, while the dipole moment of the elongated C5 dipeptide ranges between the values of the two folded C7 conformers. As expected, the antiparallel β-sheet model has a dipole moment which is very close to zero, while the helix model displays a dipole moment of ~22 D which reflects the local organization of the helix through H-bonds.

The 20 K harmonic IR signatures extracted from our CPMD simulations in the amide I−III fingerprint region are presented in Figure 2. At low temperature, each dipeptide displays specific IR signatures, and in particular folded and unfolded structures

TABLE 1: Geometric Data of the Gas Phase Alanine Peptide Conformations Obtained by Averaging over the 20 K Car−Parrinello Molecular Dynamics Performed in This Work[a]

	C7eq	C7ax	C5	β-strand	helix
	Dihedral Angles: Φ = C−N−C−C and Ψ = N−C−C−N				
Φ	−82.99 ± 2.24	72.20 ± 2.77	−150.92 ± 5.28	−144.99 ± 8.27	−74 ± 9
				−144.50 ± 9.06	(−104 ± 16)[b]
Ψ	72.89 ± 4.06	−53.82 ± 5.12	156.81 ± 6.10	148.09 ± 7.71	−13 ± 10
				147.53 ± 7.58	(−6 ± 11, +2 ± 12)[c]
	C═O···H−N H-Bond Distances				
intrastrand O···H	2.08 ± 0.05	1.93 ± 0.05	2.24 ± 0.07		2.47−2.91 ± 0.3−0.5
				2.28 ± 0.08[d]	
				2.28 ± 0.08[e]	
interstrand O···H				1.97 ± 0.06[f]	
				1.97 ± 0.07[g]	

[a] Angles are reported in degrees and distances are reported in angstroms. [b] In parentheses, we have reported the angle values obtained at the extremities of the helix chain: mean value for the C-terminal peptide group along the chain. [c] In parentheses, we have reported the angle values obtained at the extremities of the helix chain: mean values for the two N-terminal peptide groups along the chain. [d] β double strand conformer, where there are two intrastrand hydrogen bonds (between $C{=}O_{N\text{-}ter}$ and $N{-}H_{C\text{-}ter}$ groups): bottom strand. [e] β double strand conformer, where there are two intrastrand hydrogen bonds (between $C{=}O_{N\text{-}ter}$ and $N{-}H_{C\text{-}ter}$ groups): upper strand. [f] β double strand conformer, where there are two interstrand hydrogen bonds (between $C{=}O_{C\text{-}ter}$ and $N{-}H_{N\text{-}ter}$ groups). [g] β double strand conformer, where there are two interstrand hydrogen bonds (between $C{=}O_{C\text{-}ter}$ and $N{-}H_{N\text{-}ter}$ groups).

TABLE 2: Mean Values of the Dipoles of the Conformers Studied in This Work[a]

C7eq	C7ax	C5	β	helix
2.79 ± 0.19[b]	3.63 ± 0.21[b]	3.31 ± 0.32[b]	0.68 ± 0.29[b]	21.96 ± 1.43[b]
	3.49 ± 0.66[c]		1.66 ± 0.73[c]	

[a] Dipole values are reported in debyes. [b] CPMD at ∼20 K. [c] CPMD at ∼300 K. At 300 K, the two values reported in the table correspond to the dynamics where the C7ax and β-strand structures are maintained, on average; see text for explanation about the conformational dynamics of the peptides investigated here.

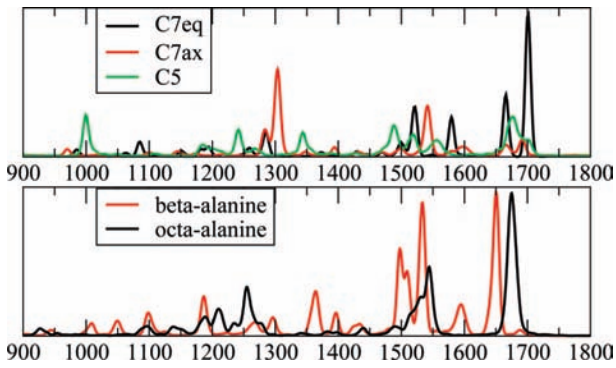

Figure 2. Harmonic infrared spectra of the three C7eq, C7ax, and C5 alanine dipeptides (top) and octaalanine and β-strand alanine peptides (bottom) calculated from Car−Parrinello molecular dynamics (CPMD) simulations at 20 K. See text for explanation. Top: C7eq (black), C7ax (red), and C5 (green). Bottom: Ala_8 (black) and β-strand (red).

give very different amide signatures, in terms of positions and relative intensities. Hence, the folded C7 conformations display two separate sharp amide I bands (C═O stretching motion), respectively located at 1702 and 1668 cm^{-1} for C7eq and 1697 and 1668 cm^{-1} for C7ax, thus identically located for both folded conformations. The highest frequency is related to the non H-bonded C═O group, while the ∼30 cm^{-1} red-shifted band is related to the H-bonded C═O group. On the contrary, the unfolded C5 has one broad band located at ∼1677 cm^{-1} where both carbonyl group stretchings arise. Overall, this broad band spans the same frequency region as the two separate amide I bands of the folded conformations, but the peak is centered between the two amide I bands of the folded C7 conformations. The 1500−1600 cm^{-1} region is composed of three bands for each dipeptide, which are systematically found red-shifted for C5. Only the 1596 cm^{-1} (C7ax), 1578 cm^{-1} (C7eq), and 1517 cm^{-1} (C5) correspond to the amide II mode (N−H bending and N−C streching motions). This active mode involves only the H-bonded N−H of the folded conformations, while both N−H participate to the IR band of the unfolded C5. This latter is 60−80 cm^{-1} red-shifted from the C7 amide II bands, which is a clear IR signature of the unfolded structure versus the folded ones. The amide III bands (N−C stretching and N−H bending) are located at 1302 and 1282 cm^{-1} for C7ax and 1282 and 1255 cm^{-1} for C7eq, each one systematically involving both amide groups, and at 1267 cm^{-1} (C-terminal amide group) and 1241 cm^{-1} (N-terminal amide group) for C5. A systematic 20−30 cm^{-1} spacing is thus obtained between both amide III modes, whatever the dipeptide structure. Note the low intensity of the amide III bands of C7eq. All other IR bands in the 1000−2000 cm^{-1} investigated domain are related to methyl and C−H group bendings and rockings. The high intensity of the 1343 cm^{-1} C_α−H rocking of the C5 conformer should be emphasized as it has no counterpart in the folded C7 structures. Note that the amide II and amide III bands of C7ax are located ∼20 cm^{-1} higher in frequency than the corresponding bands of C7eq, which could be the signature of the stronger H-bond maintaining the folded C7ax conformation. The amide II and III bands of the elongated C5 conformation are always located at lower frequency (by ∼40−80 cm^{-1}) than the corresponding bands of the folded C7 conformers. These signatures allow discrimination of folded and unfolded conformations.

All frequencies calculated here are comparable to the values obtained in the seminal and reference works of Suhai et al.,[88] Balazs,[89] and Krimm et al.,[90] and to the most recent work of Tarczay et al.[91] (scaled frequencies), with different levels of theory. Interestingly, the spectra calculated here provide a very good account of the IR spectrum of the capped alanine dipeptide recorded in rare gas matrixes at low temperature.[91,92] We take the updated spectra recorded in ref 91 in argon and krypton matrixes as references for the following discussion on the amide I and II bands, and the spectrum of ref 92 for the amide III bands (not updated in ref 91). As previously observed by Suhai et al.,[88] the matrix experimental spectra recorded at low temperature can be interpreted by the combination of the individual spectra of C7 and C5 conformers. Hence, the three amide I bands experimentally recorded can be well reproduced by the three bands arising from the two C7 conformers and the

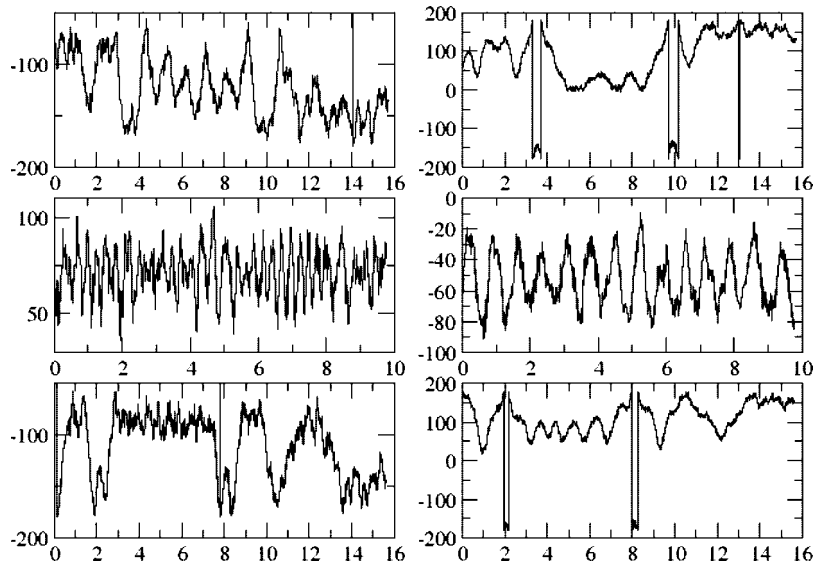

Figure 3. Evolution with time of Φ (left panel) and Ψ (right panel) dihedral angles of Ala_2 during the three 300 K Car−Parrinello MD performed in this work. The ±180° shifts of the angles are artifacts due to the definition of +180° or −180°.

C5 conformer. Though the ∼30 cm^{-1} spacing calculated between the two amide I bands of the C7 conformers is twice the experimental spacing, the 9 cm^{-1} spacing between the 1668 cm^{-1} amide I band of the C7 conformers and the 1677 cm^{-1} amide I band of the C5 conformer agrees very well with the 8 cm^{-1} spacing in the experiment. The amide II domain is less well reproduced by our calculations. The positions of the amide II bands calculated here for the C7 conformers are blue-shifted from the experiment by ∼50 cm^{-1}, while the amide II band of the C5 conformer is blue-shifted by ∼22 cm^{-1} from the third amide II band observed in the experiment. The 20 cm^{-1} frequency spacing calculated between the two amide II bands of the two C7 conformers is underestimated with respect to the experiment (∼35 cm^{-1}). As a consequence, the ∼55 cm^{-1} spacing experimentally observed between the higher and lower amide II bands is overestimated in our calculation (∼79 cm^{-1}). The use of the BLYP functional in the present work could be responsible for such deficiencies, though the same overall vibrational patterns of the amide II bands have been obtained in ref 91 with a different level of theory, with the same deficiencies in the frequency spacings in comparison to the experiment. On the contrary, the amide III bands of the three conformers calculated here give a perfect agreement with the bands recorded in ref 92, which are respectively located at 1281, 1257, and 1240 cm^{-1}.

The patterns of the 20 K harmonic IR spectrum of the 3_{10} octaalanine helix are much simpler than the ones obtained for the dipeptides in this fingerprint region. It displays only one sharp amide I band located at 1676 cm^{-1}, very close to the folded C7 lower amide I band. The amide II band is located at ∼1545 cm^{-1}, and its broadness and complex form result from strong couplings between the N−H bendings and the surrounding methyl bendings and rockings. The amide III band is located at ∼1255 cm^{-1}. Both amide II and III bands of the helix are located between the corresponding bands of the dipeptides. All other active infrared bands are due to C_α−H and methyl bendings and rockings, which give rise to relatively intense active bands. The 130 cm^{-1} spacing we obtain here between the amide I and the amide II bands is smaller than the ∼165 cm^{-1} obtained by Keiderling et al.[12] with a different theoretical approach. The form of the amide II band calculated here is very similar to the one obtained by these authors, while the amide I band we calculate clearly does not display the shoulder at higher frequency displayed in ref 12.

There are two amide I bands for the gas phase antiparallel β-strand alanine model: one of very low intensity (1688 cm^{-1}) and arising from the non H-bonded C═O groups, and one of high intensity (1650 cm^{-1}) arising from the interstrand H-bonded C═O. The spacing obtained here between the two amide I bands is comparable to the spacings calculated by Keiderling et al.[93] for various models of alanine β-strands of increasing size. Note also that the 26 cm^{-1} red shift of the intense amide I band of the β-strand alanine model from the amide I band of the helix is well-known in the liquid phase,[1] and is thus found to hold true in the gas phase. The amide II band is located at 1595 cm^{-1} and involves only the interstrand H-bonded C═O groups. It is thus blue-shifted by 50 cm^{-1} from the amide II band of the helix. We calculate two amide III bands, located at 1295 and 1263 cm^{-1}, where all amide groups participate. These bands are blue-shifted from the helix bands by ∼10−40 cm^{-1}. Again, all other active infrared bands in the domain are due to C_α−H and methyl bendings and rockings, which are particularly very intense in the 1500−1600 cm^{-1} region of the β-strand.

3.2. Dynamics and IR Spectra at 300 K. The alanine dipeptide is one of the most studied peptide models, and consequently its conformational space in the gas phase has been analyzed in detail with all kinds of ab initio theoretical methods and force fields. See, for instance, refs 74−78. Three main isomers are identified in the gas phase, referred to as the elongated C5 conformation and the folded C7eq and C7ax conformations. The energy order among these three conformations is found to depend on the theoretical level used, and whether energy differences or free energies between the isomers are taken into account. When free energies are calculated, the relative order between the three isomers also depends whether the free energies are calculated with (mostly ab initio calculations) or without (mostly umbrella sampling with classical force fields) the harmonic approximation for the entropic contribution. Therefore, C7eq and C5 are generally found to be the isomers of least energies, and their relative energy order (energy differences of ∼1−2 kcal/mol) depends on the theoretical methods.

The dynamics of the alanine dipeptide performed in this work at 300 K show a continual conformational dynamics of Ala_2,

spanning the typical C7eq and C7ax folded conformations, the unfolded C5, and mixed folded/unfolded PPII type structures. This is illustrated in Figure 3 (see also Table 1 for typical dihedral values). It is interesting to note that one trajectory remains within the C7ax basin over ~10 ps time length: see Φ and Ψ dihedrals in Figure 3, and the average value of the C7ax dipole moment at 300 K that is roughly identical to the one obtained at 20 K (Table 2). The two other dynamics show alternate periods of times during which the structure of the alanine dipeptide is arranged, on average, as an elongated C5 or PPII type, or as a folded C7eq type. Note the fluctuations of the angles around their mean values at this temperature. Dipole moments obtained over these two trajectories evolve between 0.5 and 6.0 D, reflecting the great variety of local structures of the dipeptide explored along the trajectories.

Over the overall 43 ps of dynamics performed here at ~300 K, Ala_2 thus adopts, on average, a folded conformation (C7ax or C7eq type) for 50% of the time and an unfolded conformation (C5 or PPII type) for 50% of the time. The conformational dynamics of the dipeptide between the C7eq, C5, and PPII conformations proceeds without any energy barrier at 300 K. This means that the 1−2 kcal/mol energy difference between the two conformers is lowered by entropic effects, which are taken into account in the dynamics without any approximation. This illustrates the importance of entropic contributions in structural equilibria, as already emphasized in a previous work.[38] The C7ax conformation of Ala_2 is not explored during the two trajectories displaying this conformational dynamics: C7ax is only explored during the trajectory which was initiated from the low temperature C7ax structure. Possible explanations for this are that the energy barrier to get out of the C7ax basin is higher than the energy of the 300 K dynamics performed here, or is sufficiently close to it. At best, the crossing of the barrier is a rare event that is difficult to probe during the short dynamics performed in this work. We have not investigated this further, as this was not our main purpose.

Taking into account the conformational dynamics of the dipeptide at 300 K, our calculated IR spectrum of Ala_2 has therefore been averaged over the three dynamics for an overall 43 ps of dynamics. The dynamics which remains within the C7ax basin over 10 ps gives us the opportunity to extract the 300 K IR signatures of the folded C7ax conformation separately, and to directly compare it to its 20 K harmonic counterparts. In that case, as can be seen from Figures 2 and 4, temperature broadens the active bands of C7ax and slightly shifts the positions of the bands.

Interestingly, the 3_{10} octaalanine helix skeleton is not maintained over the 13 ps trajectory at 300 K. In less than 6 ps, the peptide evolves toward a stable elongated chain maintained through $i \rightarrow i+1$ H-bonds between two successive N−H and C═O peptide groups, as illustrated in Figure 1f. The average H-bond distances range between 2.0 and 2.6 Å, depending on the position along the chain. Disappearance of the helix chain of the neutral alanine peptide at room temperature is nicely in agreement with the results of Jarrold et al.[57] and Dugourd et al.[58] Average values of Φ and Ψ angles over the chain are −80°/80°, which are very close to the values obtained for the typical folded C7eq dipeptide structure. The chain can thus be seen as organized as successive C7eq peptide groups. We do not claim here that this is the final structure of the Ala_8 peptide; longer dynamics would be needed for that. This relatively short dynamics has the main purpose of demonstrating that the helix is lost, and that the peptide unfolds at 300 K. The IR spectrum of the octaalanine peptide presented in Figure 4 has been calculated over the last 7 ps of dynamics where the elongated structure is maintained. Strikingly, exactly the same spectral features are obtained over the first part of the dynamics where the 3_{10} helix is progressively unfolding.

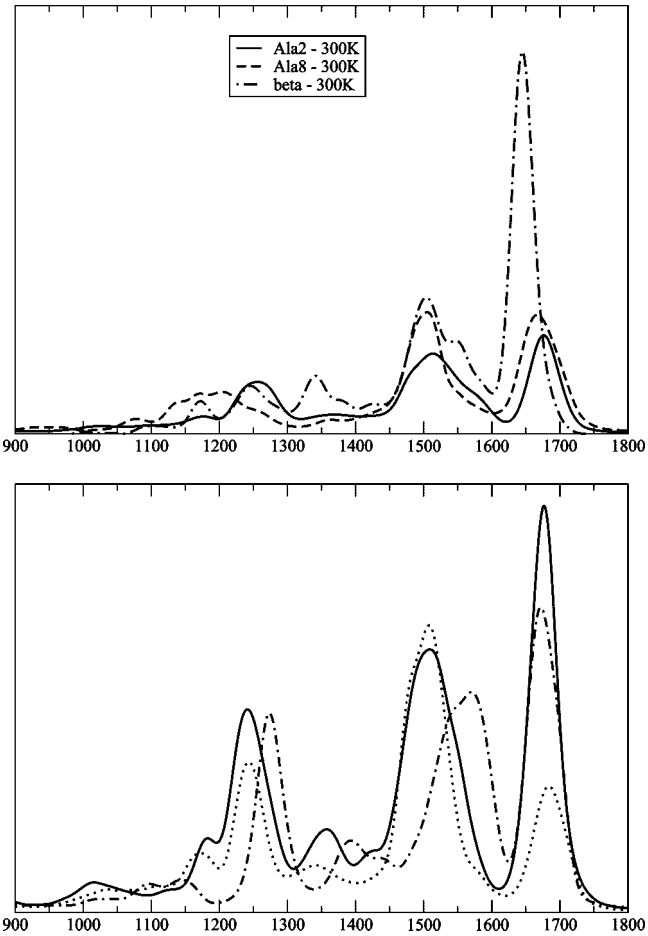

Figure 4. Top: Infrared spectra of alanine dipeptide Ala_2 (solid line), octaalanine peptide Ala_8 (dashed line), and β-strand alanine peptide (dot−dashed line) calculated from Car−Parrinello molecular dynamics (CPMD) simulations at 300 K. Bottom: Highlights on the alanine dipeptide IR spectra generated from the three separate 300 K trajectories. Dot−dashed line: IR spectrum of Ala_2 from the 300 K dynamics during which the C7ax conformation is solely explored. Solid line and dotted line: IR spectra of Ala_2 from the 300 K dynamics displaying a conformational dynamics of the peptide. See text for more explanation.

On the contrary, the β-strand alanine model is maintained during the 300 K dynamics, though slightly distorted from the 20 K structure. Hence, the C═O···H−N interstrand H-bonds are elongated to an average 2.3 Å, but they are energetically strong enough for the interstrand to be maintained. We find that the average dipole of the β-strand evolves from ~0.7 ± 0.3 D at 20 K to ~1.7 ± 0.7 D at 300 K. The relatively small fluctuations of the dipole at 300 K reveal the exploration of one single average conformation only during the dynamics. The increase of the mean dipole moment at 300 K is the result of the slight distortion in the antiparallel organization of the two strands.

The 300 K IR spectra of the dipeptide, the β-strand, and the octapeptide calculated from our CPMD trajectories are superimposed in Figure 4. Separate contributions from the three MD Ala_2 trajectories are also reported at the bottom of Figure 4. We immediately observe that the amide I band is identically located at ~1670−1680 cm^{-1} for Ala_2 and Ala_8 at room

temperature. Band positions differ by 10 cm^{-1} for the three Ala_2 dynamics, which is just within the 5−10 cm^{-1} expected theoretical error. It is really remarkable that the amide I bands of Ala_2 and Ala_8 are identically located, even though the local structures of the two peptides, i.e., H-bonds and dihedral angles of the chains, are not comparable on average. At 300 K, the Ala_8 peptide is elongated and is organized as a succession of C7eq dipeptides along the chain, while the conformational dynamics displayed by Ala_2 at 300 K gives an equal mixture of folded conformations (C7eq and C7ax types) and unfolded conformations (C5 and PPII types), whose signatures are all contained within the amide I band. Thus, though the C═O groups of Ala_2 and Ala_8 chains are involved in different hydrogen bonding patterns along the time, their vibrational contributions to the amide I band are identical. Note also that the amide I band appears as slightly asymmetrical in both Ala_2 and Ala_8, which suggests that there are, on average, contributions from symmetric and antisymetric vibrations of the C═O groups of the chains.[94] We observe that the 300 K amide I band of the β-strand peptide (∼1645 cm^{-1}) is almost nonshifted from its 20 K counterpart (1650 cm^{-1}), and is ∼30 cm^{-1} red-shifted from the amide I band of Ala_2 and Ala_8. This shift is substantial enough to be detected in experiments. We therefore find that the amide I band signature does not allow distinguishing the structures displayed by Ala_2 and Ala_8, but gives a clear opportunity to distinguish these "not so well organized" structures from the "more organized" intermolecular H-bonded β-strand peptide, at room temperature and in the gas phase.

The amide II bands of the dipeptide, octapeptide, and β-strand peptide are identically peaked at ∼1510 cm^{-1}, but the band shapes and broadness are different. The amide II band of Ala_2 and the β-strand are broad and asymmetric due to the shoulder located at ∼1560−1570 cm^{-1}, while the amide II band of Ala_8 is globally symmetric with no appearance of the shoulder at higher frequency. The broadening and shoulder of the amide II band of Ala_2 are due to the folded C7ax conformation: the spectrum extracted from the trajectory of C7ax at 300 K indeed displays an amide II signature that is blue-shifted by ∼50 cm^{-1} from the other amide II bands of Ala_2 (bottom of Figure 4). This blue shift is entirely due to the N-terminal amide N−H group of C7ax which is hydrogen bonded to the C-terminal C═O group of the peptide along the 300 K trajectory. The main peak of the amide II band located at ∼1510 cm^{-1} is due to the bending motions of all N−H groups of the dipeptide, and involves only the C-terminal N−H that is not involved in hydrogen bonding along the trajectory of the folded C7ax. Following the same patterns, the main amide II peak of the β-strand is related to the N−H groups that are not involved in H-bonds along the 300 K trajectory, and the blue-shifted shoulder is related to the N−H groups involved in the interstrand H-bonds. These blue shifts are consistent with the formation of H-bonds, as previously observed.[53,54] We thus find that, at 300 K, the amide II band is identically located at ∼1510 cm^{-1} for all the alanine peptides investigated here, and the broadness of this band displayed by a shoulder that is blue-shifted by ∼50 cm^{-1} from the main peak is the signature of H-bonds in the peptide conformations. Remarkably, the spacing between amide I and amide II is ∼163 cm^{-1} for Ala_2 and Ala_8, and is reduced to ∼143 cm^{-1} for the β-strand alanine peptide, which can be used as a way of discriminating a β-strand structure from other peptide structures at 300 K in the gas phase.

The domain of the amide III vibrational signatures extends between ∼1100 and ∼1300 cm^{-1} for the alanine peptides investigated here at 300 K in the gas phase. The broad band located at ∼1180 cm^{-1} comes from the elongated octaalanine peptide, and therefore displays a large spacing of ∼400 cm^{-1} from the amide II band. Note that the Ala_8 amide III band has been red-shifted by 75 cm^{-1} in going from 20 K (3_{10} helix) to 300 K (elongated structure). There are two amide III bands associated with the β-strand model at ∼1175 and ∼1245 cm^{-1}, which are respectively related to the non-H-bonded and H-bonded peptide groups of the strands. The two signatures are distinguishable from each other (no overlap between the two bands), which provides a way of discriminating the H-bonded stranded peptide. We find one amide III band for the dipeptide, less broad than the one of the octapeptide, which is located at ∼1257 cm^{-1}, almost identically to one of the amide III bands of the β-strand. Note again that the folded C7ax amide III signature is blue-shifted from the average amide III signature of Ala_2, as it is located at ∼1270 cm^{-1}, and therefore gives rise to the broadness of the final Ala_2 amide III band. The amide II−amide III spacing of the dipeptide is ∼80 cm^{-1} smaller than the one from the octapeptide. The amide III band thus appears as a good vibrational signature of the alanine peptide structure. The elongated octapeptide displays the broadest band that is located at the lowest frequency, while the H-bonded β-strand displays two nonoverlapping amide III bands separated by ∼70 cm^{-1}. The dipeptide has also only one broad amide III band which is blue-shifted by ∼70−80 cm^{-1} from the amide III signature of the octapeptide.

The low intensity bands arising between 1300 and 1400 cm^{-1} are due to the bending and out-of-plane movements of the C_α−H groups of the peptides, which happen to be more active for the H-bonded β-strand peptide.

4. Discussion and Conclusions

The 300 K trajectories have shown that the neutral Ala_2 alanine dipeptide undergoes a continuous conformational dynamics at that temperature, and that the neutral Ala_8 polypeptide does not maintain a helical structure and unfolds toward an elongated structure. The trajectories of the dipeptide have shown that Ala_2 conformations can be viewed as an equal mixture of unfolded (C5 and PPII types) and folded (C7eq and C7ax types) structures, and the conformational dynamics and unfoldings proceed without an energy barrier at 300 K. On the contrary, the β-strand structure is maintained at room temperature, through long interstrand H-bonds which are energetically sufficient to maintain the two strands together. The conformational results obtained for the helix and β-strand are nicely in agreement with the mobility measurements of Jarrold's group,[57] confirming in a direct way that gas phase neutral helices are destabilized at room temperature while gas phase β-strands are energetically more stable.

One question we want to answer in the present work is whether the direct inclusion of temperature for the calculation of the IR spectra through molecular dynamics is mandatory for the comprehension of the vibrational signatures of neutral peptides in the gas phase. In other words, could a harmonic calculation of the spectra, subsequently corrected by a Boltzmann factor for the different isomers, be sufficient for the understanding of the vibrational patterns, and would it be sufficient for the interpretation of experimental spectra recorded at finite temperature? This is the standard way applied in the literature. In the present work, we have carried out MD simulations at low temperature and room temperature in order to calculate the infrared spectra of alanine peptide models at these two temperatures, and answer that question in a direct way. On the one hand, the low temperature 20 K dynamics is

indeed the way to calculate harmonic infrared spectra that are identical to the harmonic spectra calculated through the Hessian diagonalization in standard ab initio calculations. On the other hand, the 300 K dynamics allows exploration of the conformational dynamics of the peptides (though obviously depending on the energy barriers between the isomers) and opens the way to account for vibrational couplings and vibrational anharmonicities in the infrared spectra, without approximations.

As the β-strand structure is maintained at 300 K, we can compare the harmonic spectrum of the strand calculated at 20 K with the nonharmonic one at 300 K, and figure out the effects of vibrational couplings and anharmonicities on the band positions and band shapes. Temperature and dynamics are mainly responsible for band broadenings, and vibrational couplings and anharmonicities should result in displacements of the band positions. The amide I band of the strand has been red-shifted by 6 cm^{-1} in going from 20 to 300 K, which measures the small anharmonicities of the C═O vibrational stretches at 300 K. The 300 K amide II band of the strand is modified from the 20 K amide II band. In the harmonic spectrum, there is indeed one active band located at $\sim$1595 cm^{-1} that is solely due to the H-bonded N−H groups. At 300 K, the amide II band has been broadened, and especially red-shifted. The main peak is indeed located at $\sim$1510 cm^{-1}, and the band broadness comes from the shoulder located at $\sim$1560−1570 cm^{-1}. Interestingly, vibrations of the H-bonded N−H groups arise only within the band shoulder at higher frequency, thus red-shifted by $\sim$25−35 cm^{-1} from the harmonic vibrations, while the non-H-bonded N−H groups give rise to the active main peak at $\sim$1510 cm^{-1}, which was not IR active in the harmonic spectrum. The discrimination of H-bonded and non-H-bonded amide groups is also present in the amide III bands obtained at 300 K, contrary to the harmonic spectrum. Hence, the two amide III bands of the harmonic spectrum (1295 and 1263 cm^{-1}) both involved the vibrations of all peptide groups, while the $\sim$1245 cm^{-1} band at 300 K is only due to the vibrations of the H-bonded peptide groups and the $\sim$1175 cm^{-1} band to the vibrations of the non-H-bonded peptide groups. We thus obtain a $\sim$70 cm^{-1} blue shift of the vibrations of the H-bonded groups with respect to the non-H-bonded peptide groups. The amide III domain is therefore especially modified from its harmonic part. The $\sim$30 cm^{-1} spacing between the two harmonic amide III bands has been increased to $\sim$70 cm^{-1}, splitting contributions from H-bonded and non-H-bonded peptide groups. This should be the result of anharmonic effects on the N−H and N−C stretching vibrational motions.

There is therefore no way that the harmonic spectrum calculated for the β-strand could be used in order to interpret precisely the vibrational features at 300 K. The shifts of the positions of the amide II and amide III bands obtained at 300 K would not be accounted for, nor would the splitting between the two amide II bands or between the two amide III bands.

The situation is even more critical for the octaalanine peptide, as the structure evolves from a well-organized 3_{10} helix at low temperature to an elongated peptide at 300 K. Therefore, the differences between the harmonic 20 K spectrum of the helix and the 300 K spectrum of the elongated peptide contain the signatures of both the conformational transformation and the vibrational anharmonicities of the modes. These are unfortunately impossible to separate from the conformational transformation signatures, from the simple comparison between the two spectra we make here. Strikingly, we find that the amide I band is not sensitive to the local conformation of the peptide, as its position is barely modified from 20 to 300 K (6 cm^{-1} blue shift), despite the huge conformational modification of the octapeptide. The 300 K amide II band has been red-shifted by $\sim$35 cm^{-1} from the harmonic spectrum, and its shape has been simplified. The red shift results from the loss of hydrogen bonds of the amide N−H groups in the elongated octapeptide structure. The huge $\sim$75 cm^{-1} red shift of the amide III at 300 K reveals the loss of the helix conformation toward the elongated conformation of the peptide. The harmonic spectrum of the helix cannot therefore be used to get the IR signatures of the peptide at higher temperature.

The conformational dynamics of Ala_2 at 300 K makes the interpretation of the IR spectrum of the peptide interesting and more challenging. The dynamics have indeed shown that Ala_2 conformations can be viewed as an equal mixture of unfolded (C5 and PPII types) and folded (C7eq and C7ax types) structures. As a natural question, could we therefore take the harmonic spectra of these individual conformations, apply the Boltzmann factors (calculated in the harmonic approximation) to these spectra, and recover a final spectrum that would be comparable to the one we have calculated at 300 K? Note here the ambiguity of this method as, depending on the theoretical methods employed, either C7eq or C5 is found as the conformer of least energy with $\sim$1−2 kcal/mol energy difference.[74−78] Such an energy difference gives a Boltzmann factor of the conformer of higher energy in the range 0.1−0.4, leading to a relatively low contribution of this conformer in the final vibrational spectrum. Moreover, this method does not account for the PPII structures nor all other conformations which are indeed explored during the dynamics, and which contribute to the final IR spectrum.

We have seen that the amide I band of Ala_2 at 300 K is located at $\sim$1680 cm^{-1}, which is basically representative of the position of the amide I band of the elongated C5 conformer in the harmonic spectrum. The 300 K amide II band, which is composed of a main peak located at $\sim$1510 cm^{-1} and of a shoulder at $\sim$1560−1570 cm^{-1}, could be viewed as a combination of the harmonic amide II band of the elongated C5 conformer ($\sim$1517 cm^{-1}) and of the harmonic amide II band of the folded C7 conformers ($\sim$1578 and $\sim$1596 cm^{-1}), with these latter ones being red-shifted by anharmonic vibrational effects. Though the broad 300 K amide III band that is located between 1230 and 1290 cm^{-1} globally contains all the harmonic amide III signatures of the folded and elongated conformers, its main position at $\sim$1257 cm^{-1} would be missed when considering only the harmonic spectra in order to get the spectrum at room temperature.

The present ab initio MD simulations therefore illustrate the pivotal role of temperature on gas phase polypeptide structures, dynamics, and IR signatures, and how MD simulations can help unravel these properties at room temperature in the gas phase, in a direct way. This should be taken into account when interpreting experimental infrared spectra acquired at room temperature. The structures and IR spectra of the gas phase neutral peptides obtained from geometry optimizations and harmonic vibrational analysis should not be used blindly for interpreting the experiments performed at room temperature.

The other question we want to answer to is to what extent the amide I, II, and III domain is sensitive enough to the local structure of the peptides, to provide vibrational signatures that can be definitely used to assess the peptide conformation at 300 K. We have seen that the 20 K harmonic IR spectra give very specific amide I−III vibrational signatures which in turn allow discrimination of folded and unfolded conformational isomers of the peptide. We have shown that this is not true at 300 K

anymore. We have shown that the position of the amide I band only allows discrimination of the intermolecular H-bonded β-strand peptide. There is indeed a systematic $\sim$30 cm^{-1} spacing between the amide I band of the β-strand peptide and the amide I band of the other peptides, whatever the temperature. There are thus no anharmonic effects on the amide I bands, as this spacing is maintained from harmonic (20 K) to nonharmonic (300 K) dynamics.

Interestingly, the complex harmonic band shape of the amide II band of the octaalanine peptide is lost in going from the 3_{10} H-bonded structure at 20 K to the elongated non-H-bonded structure at 300 K. At 300 K, we indeed obtain a very simple and symmetrical amide II band which clearly reflects that the H-bonds have been lost as well as the couplings of the N$-$H bendings with the surrounding C$-$H bendings. Moreover, the amide II band is downshifted by $\sim$35 cm^{-1} at 300 K. On the contrary, the amide II band shape becomes more complex for the β-strand peptide at 300 K, where the different signatures of the H-bonded and non-H-bonded N$-$H give rise to the broadness and shoulder of the band. In fact, only the H-bonded N$-$H groups gave an active amide II signature in the harmonic spectrum (1595 cm^{-1}), while both H-bonded and non-H-bonded amide groups give rise to the broad and active amide II band at 300 K. The higher frequency part of this band is therefore downshifted by $\sim$20 cm^{-1} from the harmonic band, while the non-H-bonded amide groups now give rise to an intensity of the related band at $\sim$1510 cm^{-1}. This main band is now located identically to the Ala_2 peptide at 300 K.

The amide III band of Ala_2 obtained at 300 K globally includes the harmonic amide III signatures of all folded and unfolded conformers, which explains its broadness. On the contrary, the amide III band of Ala_8 obtained at 300 K is highly red-shifted from the 20 K band by $\sim$70 cm^{-1}, which arises from the conformational transformation of the octapeptide from an helix to an elongated structure. The amide III patterns of the β-strand are also modified between 20 and 300 K, as the spacing between the two bands is noticeably increased ($\sim$32 cm^{-1} at 20 K versus $\sim$70 cm^{-1} at 300 K). All amide groups participate in the two amide III bands at 20 K, while the two amide III bands at 300 K have separate contributions from the H-bonded (1245 cm^{-1}) and non-H-bonded (1175 cm^{-1}) peptide groups of the two strands.

At room temperature, the amide I band only allows discrimination of β-strands from other conformations, using the 26 cm^{-1} red shift of the amide I of the β-strand with respect to the other amide I bands. The amide II band is identically located at $\sim$1510 cm^{-1} for all structures, but its broadness and band shape is an indicator that the vibrations of the N$-$H groups are split depending whether these groups are involved or not in hydrogen bonding. Anharmonicities and vibrational couplings of the intermolecular H-bonds are therefore responsible for the splittings. The amide III domain is a good probe of the local strcuture of the alanine peptides. The β-strand displays two nonoverlapping amide III bands, and is the only conformer with two such active bands. Ala_2 and Ala_8 peptides have one broad amide III band which is upshifted by $\sim$70 cm^{-1} for the dipeptide.

We also find that the active bands arising from the bendings and rockings of $C_\alpha$$-$H and methyl C$-$H groups at 20 K do not have any activity at 300 K, apart from the β-strand peptide where the 1300$-$1400 cm^{-1} band is still active.

Our results therefore show that one has to be cautious about the interpretation of infrared spectra recorded at room temperature based on standard harmonic calculations of IR spectra, as the vibrational patterns from these calculations could be misleading. The usual application of a Boltzmann factor reweighting the final IR spectrum from the individual spectra of each identified isomer would broaden the initial (0 K) vibrational patterns. Band broadening could then, to a certain extent, be comparable to the MD results. Nonetheless, all anharmonic vibrational effects which are taken into account in a direct way in the MD simulation of the spectra would inevitably be missed, and subsequent band displacements and band splittings would not be obtained. The harmonic vibrational signatures lack the conformational dynamics information that is pivotal for a precise interpretation of the IR signatures which are experimentally recorded at room temperature. This should be taken into account when seeking the IR signatures of polypeptide helices and strands in gas phase room temperature experiments. We hope the present simulations will convince and encourage the community to go beyond standard harmonic structural and vibrational calculations in order to interpret linear IR (1D-IR) gas phase room temperature experiments on polypeptides.

The present calculations have also shown that IR fingerprints at 300 K are not easily related to the local structure of the peptides. Amide I and amide II bands cannot be used as precise structural fingerprints, as elongated and folded structures globally share the same patterns. These two bands can only serve as indicators of H-bonded peptide groups whithin the chains. On the contrary, the amide III band can act as a structural probe. Depending on the organization of the peptide chain (Φ and Ψ dihedral angles), the amide III band is scattered over an $\sim$300 cm^{-1} range. The number of amide III bands is another structural probe, as the H-bonded β-strand peptide possesses two nonoverlapping bands while the elongated and folded single strand peptide possesses one broad band.

Acknowledgment. M.-P.G. thanks IDRIS (Orsay, France) for generous access to their computational facilities and acknowledges support from Genopole-France through the program "ATIGE" Action Thématique Incitative de Génopole.

References and Notes

(1) Krimm, S.; Bandekar, J. *Adv. Protein 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) Herrmann, C.; Neugebauer, J.; Reiher, M. *New J. Chem.* **2007**, *31*, 818.
(6) Taillandier, E.; Peticolas, W.; S. Adam, T. H.-D.; Igolen, J. *Spectrochim. Acta. A* **1990**, *46*, 107.
(7) Powell, J.; Peticolas, W.; Genzel, L. *J. Mol. Struct.* **1991**, *247*, 107.
(8) Fritzche, H. *J. Mol. Struct.* **1991**, *242*, 245.
(9) Semenov, M.; Bolbukh, T.; Mallev, V. *J. Mol. Struct.* **1997**, *408*$-$*409*, 213.
(10) Huang, R.; Kubelka, J.; Barber-Armstrong, W.; Silva, R.; Decatur, S.; Keiderling, T. *J. Am. Chem. Soc.* **2004**, *126*, 2346.
(11) Schweitzer-Stenner, R.; Eker, F.; Huang, Q.; Griebenov, K. *J. Am. Chem. Soc.* **2001**, *123*, 9628.
(12) Bour, P.; Kubelka, J.; Keiderling, T. *Biopolymers* **2002**, *65*, 45.
(13) Singh, B. R. Infrared analysis of peptides and proteins: Principles and Applications. In *ACS Symposium Series*; American Chemical Society: Washington, DC, 2000.
(14) Schermann, J. P. *Spectroscopy and Modeling of Biomolecular Building Blocks*; Elsevier Science: New York, 2007.
(15) Stearns, J. A.; Boyarkin, O. V.; Rizzo, T. R. *J. Am. Chem. Soc.* **2007**, *129*, 13820.
(16) Stearns, J. A.; Mercier, S.; Seaiby, C.; Guidi, M.; Boyarkin, O. V.; Rizzo, T. R. *J. Am. Chem. Soc.* **2007**, *129*, 11814.
(17) Brenner, V.; Piuzzi, F.; Dimicoli, I.; Tardivel, B.; Mons, M. *J. Phys. Chem. A* **2007**, *111*, 7347.

(18) Rijs, A. M.; Crews, B. O.; de Vries, M. S.; Hannam, J. S.; Leigh, D. A.; Fanti, M.; Zerbetto, F.; Buma, W. J. *Angew. Chem., Int. Ed.* **2008**, *47*, 3174.
(19) Carnegie, P. D.; Bandyopadhyay, B.; Duncan, M. A. *J. Phys. Chem. A* **2008**, *112*, 6237.
(20) Douberly, G. E.; Ricks, A. M.; Ticknor, B. W.; Mckee, W. C.; Schleyer, P. v. R.; Duncan, M. A. *J. Phys. Chem. A* **2008**, *112*, 1897.
(21) Kolaski, M.; Lee, H. M.; Choi, Y. C.; Kim, K. S.; Tarakeshwar, P.; Miller, D. J.; Lisy, J. M. *J. Chem. Phys.* **2007**, *126*, 074302.
(22) Loh, Z. M.; Wilson, R. L.; Wild, D. A.; Bieske, E. J.; Lisy, J. M.; Njegic, N. B.; Gordon, M. S. *J. Phys. Chem. A* **2006**, *110*, 13736.
(23) Lemaire, J.; Boissel, P.; Heninger, M.; Mauclaire, G.; Bellec, G.; Mestdagh, H.; Simon, A.; Caer, S. L.; Ortega, J.; Glotin, F.; Maitre, P. *Phys. Rev. Lett.* **2002**, *89*, 273002-1, and references therein.
(24) Oomens, J.; Meijer, G.; Helden, G. *J. Phys. Chem. A* **2001**, *105*, 8302.
(25) Kapota, C.; Lemaire, J.; Maitre, P.; Ohanessian, G. *J. Am. Chem. Soc.* **2004**, *126*, 1836.
(26) Correia, C. F.; Balaj, P. O.; Scuderi, D.; Maitre, P.; Ohanessian, G. *J. Am. Chem. Soc.* **2008**, *130*, 3359.
(27) Vaden, T. D.; de Boer, T. S. J. A.; Simons, J. P.; Snoek, L. C. *Phys. Chem. Chem. Phys.* **2008**, *10*, 1443.
(28) Bush, M. F.; Oomens, J.; Saykally, R. J.; Williams, E. R. *J. Am. Chem. Soc.* **2008**, *130*, 6463.
(29) Kapota, C.; Ohanessian, G. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3744.
(30) Lucas, B.; Gregoire, G.; Lemaire, J.; Maitre, P.; Ortega, J.; Rupenyan, A.; Reimann, B.; Schermann, J.; Desfrancois, C. *Phys. Chem. Chem. Phys.* **2004**, *6*, 2659.
(31) Lucas, B.; Gregoire, G.; Lemaire, J.; Maitre, P.; Glotin, F.; Schermann, J.; Desfrancois, C. *Int. J. Mass Spectrom.* **2005**, *243*, 105.
(32) Chin, W.; Piuzzi, F.; Dognon, J.; Dimicoli, I.; Tardivel, B.; Mons, M. *J. Am. Chem. Soc.* **2005**, *127*, 11900.
(33) Salpin, J.; Guillaumont, S.; Tortajada, J.; MacAleese, L.; Lemaire, J.; Maitre, P. *ChemPhysChem* **2006**, *8*, 2235.
(34) Vaden, T. D.; de Boer, T. S. J. A.; Simons, J. P.; Snoek, L. C.; Suhai, S.; Paizs, B. *J. Phys. Chem. A* **2008**, *112*, 4608.
(35) Grégoire, G.; Gaigeot, M.; Marinica, D.; Lemaire, J.; Schermann, J.; Desfrançois, C. *Phys. Chem. Chem. Phys.* **2007**, *9*, 3082.
(36) Gregurick, S. K.; Chaban, G. M.; Gerber, B. *Biophys. J.* **2001**, *80*, 303.
(37) Barone, V. *J. Chem. Phys.* **2005**, *122*, 014108.
(38) Marinica, C.; Grégoire, G.; Desfrançois, C.; Schermann, J.; Borgis, D.; Gaigeot, M.-P. *J. Phys. Chem. A* **2006**, *110*, 8802.
(39) Allen, M.; Tildesley, D. *Computer simulation of liquids*; Oxford Science Publications: New York, 1997.
(40) Frenkel, D.; Smit, B. *Understanding molecular simulations*, 2nd ed.; Academic Press: New York, 2002.
(41) McQuarrie, D. *Statistical Mechanics*; Harper-Collins Publishers: New York, 1976.
(42) Vuilleumier, R.; Borgis, D. *J. Chem. Phys.* **1999**, *111*, 4251.
(43) Tassaing, T.; Danten, Y.; Besnard, M.; Zoidis, E.; Yarwood, J.; Guissani, Y.; Guillot, B. *Mol. Phys.* **1995**, *84*, 769.
(44) Guillot, B. *J. Chem. Phys.* **1991**, *95*, 1543.
(45) Glover, W.; Madden, P. *J. Chem. Phys.* **2004**, *121*, 7293.
(46) Pavlatou, E.; Madden, P.; Wilson, M. *J. Chem. Phys.* **1997**, *107*, 10446.
(47) Madden, P.; Impey, R. *Chem. Phys. Lett.* **1986**, *123*, 502.
(48) Marzari, N.; Vanderbilt, D. *Phys. Rev. B* **1997**, *56*, 12847–12865.
(49) Car, R.; Parrinello, M. *Phys. Rev. Lett.* **1985**, *55*, 2471.
(50) Silvestrelli, P.; Parrinello, M. *J. Chem. Phys.* **1999**, *111*, 3572.
(51) Silvestrelli, P.; Bernasconi, M.; Parrinello, M. *Chem. Phys. Lett.* **1997**, *277*, 478.
(52) Bernasconi, M.; Silvestrelli, P.; Parrinello, M. *Phys. Rev. Lett.* **1998**, *81*, 1235.
(53) Gaigeot, M.-P.; Sprik, M. *J. Phys. Chem. B* **2003**, *107*, 10344.
(54) Gaigeot, M.-P.; Vuilleumier, R.; Sprik, M.; Borgis, D. *J. Chem. Theory Comput.* **2005**, *1*, 772.
(55) Gaigeot, M.; Martinez, M.; Vuilleumier, R. *Mol. Phys.* **2007**, *105*, 2857.
(56) Barone, V.; Cimino, P.; Crescenzi, O.; Pavone, M. *J. Mol. Struct. (THOECHEM)* **2007**, *811*, 323.
(57) Jarrold, M. *Phys. Chem. Chem. Phys.* **2007**, *9*, 1659.
(58) Poulain, P.; Calvo, F.; Antoine, R.; Broyer, M.; Dugourd, P. *Eur. Phys. Lett.* **2007**, *79*, 66003.
(59) Joly, L.; Antoine, R.; Allouche, A. R.; Broyer, M.; Lemoine, J.; Dugourd, P. *J. Am. Chem. Soc.* **2007**, *129*, 8428.
(60) Compagnon, I.; Oomens, J.; Meijer, G.; von Helden, G. *J. Am. Chem. Soc.* **2006**, *128*, 3592.
(61) Besley, N. A. *Philos. Trans. R. Soc. A* **2007**, *365*, 2799.
(62) Kubelka, J.; Silva, R. G. D.; Keiderling, T. *J. Am. Chem. Soc.* **2001**, *124*, 5325.
(63) Silva, R. G. D.; Kubelka, J.; Bour, P.; Decatur, S.; Keiderling, T. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 8318.
(64) Huang, R.; Kubelka, J.; Barber-Armstrong, W.; Silva, R.; Decatur, S.; Keiderling, T. *J. Am. Chem. Soc.* **2004**, *126*, 2346.
(65) Marx, D.; Tuckerman, M. E.; Parrinello, M. *J. Phys.: Condens. Matter A* **2000**, *12*, 153–159.
(66) Gaigeot, M.-P.; Sprik, M. *J. Phys. Chem. B* **2004**, *108*, 7458.
(67) Trouillier, N.; Martins, J. *Phys. Rev. B* **1991**, *43*, 1993.
(68) Kleinman, L.; Bylander, D. *Phys. Rev. Lett.* **1982**, *48*, 1425.
(69) Becke, A. *Phys. Rev. A* **1988**, *38*, 3098.
(70) Lee, C.; Yang, W.; Parr, R. *Phys. Rev. B* **1988**, *37*, 785.
(71) CPMD: copyright IBM Corp. 1990−2006; copyright MPI für Festkorperforschung Stuttgart 1997−2001.
(72) Spezia, R.; Tournois, G.; Tortajada, J.; Cartailler, T.; Gaigeot, M.-P. *Phys. Chem. Chem. Phys.* **2006**, *8*, 2040.
(73) Martyna, G.; Tuckerman, M. *J. Chem. Phys.* **1999**, *110*, 2810.
(74) Roterman, I. K.; Lamberg, M. H.; Gibson, K. D.; Scheraga, H. A. *J. Mol. Biol. Struct. Dyn.* **1989**, *7*, 421.
(75) Tobias, D. J.; Brooks, C. L. *J. Phys. Chem.* **1992**, *96*, 3864.
(76) Vargas, R.; Garza, J.; Hay, B. P.; Dixon, D. A. *J. Phys. Chem. A* **2002**, *106*, 3213.
(77) Feig, M.; MacKerell, A. D.; Brooks, C. L. *J. Phys. Chem. B* **2003**, *107*, 2831.
(78) Rosso, L.; Abrams, J. B.; Tuckerman, M. E. *J. Phys. Chem. B* **2005**, *109*, 4162.
(79) King-Smith, R.; Vanderbilt, D. *Phys. Rev. B* **1993**, *47*, 1651.
(80) Vanderbilt, D.; King-Smith, R. *Phys. Rev. B* **1993**, *48*, 4442.
(81) Iftimie, R.; Tuckerman, M. *J. Chem. Phys.* **2005**, *122*, 214508.
(82) Borysow, J.; Moraldi, M.; Frommhold, L. *Mol. Phys.* **1985**, *56*, 913.
(83) Ramirez, R.; Lopez-Ciudad, T.; Kumar, P.; Marx, D. *J. Chem. Phys.* **2004**, *121*, 3973.
(84) Tangney, P. *J. Chem. Phys.* **2006**, *124*, 044111.
(85) Tangney, P.; Scandolo, S. *J. Chem. Phys.* **2002**, *116*, 14.
(86) Todorova, T.; Seitsonen, A. P.; Hutter, J.; Kuo, I. F. W.; Mundy, C. J. *J. Phys. Chem. B* **2006**, *110*, 3685.
(87) Guidon, M.; Schiffmann, F.; Hutter, J.; VandeVondele, J. *J. Chem. Phys.* **2008**, *128*, 214104.
(88) Jalkanen, K. J.; Suhai, S. *Chem. Phys.* **1996**, *208*, 81.
(89) Balazs, A. *J. Phys. Chem.* **1990**, *94*, 2754.
(90) Cheam, T.; Krimm, S. *J. Mol. Struct. (THOECHEM)* **1989**, *188*, 15.
(91) Pohl, G.; Perczel, A.; Vass, E.; Magyarfalvi, G.; Tarczay, G. *Phys. Chem. Chem. Phys.* **2007**, *9*, 4698.
(92) Grenie, Y.; Avignon, M.; Garrigou-Lagrange, C. *J. Mol. Struct.* **1975**, *24*, 293.
(93) Kubelka, J.; Keiderling, T. *J. Am. Chem. Soc.* **2001**, *123*, 6142.
(94) Kwac, K.; Lee, K. K.; Han, J. B.; Oh, K. I.; Cho, M. *J. Chem. Phys.* **2008**, *128*, 105106.

JP807550J