JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

Finite Temperature Infrared Spectra from Polarizable Molecular Dynamics Simulations

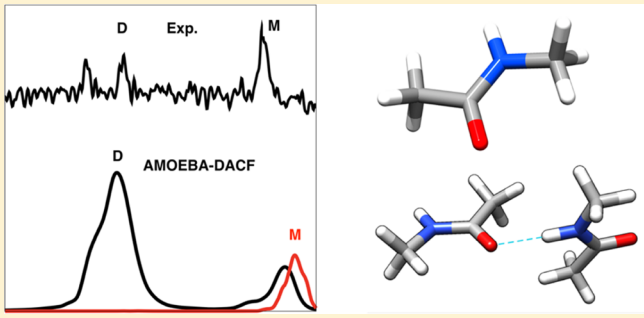
David Semrouni,[†,¶] Ashwani Sharma,[†] Jean-Pierre Dognon,[‡] Gilles Ohanessian,*[,†] and Carine Clavaguéra*[,†]

[†]Laboratoire de Chimie Moléculaire, Ecole polytechnique, CNRS, 91128 Palaiseau Cedex, France, and

[‡]CEA/Saclay, DSM/IRAMIS/NIMBE, CNRS, Laboratoire de Chimie Moléculaire et de Catalyse pour l'Energie, 91191 Gif-sur-Yvette, France

Supporting Information

ABSTRACT: Infrared spectra of biomolecules are obtained from molecular dynamics simulations at finite temperature using the AMOEBA force field. Diverse examples are presented such as *N*-methylacetamide and its derivatives and a helical peptide. The computed spectra from polarizable molecular dynamics are compared in each case to experimental ones at various temperatures. The role of high-level electrostatic treatment and explicit polarization, including parameters improvements, is highlighted for obtaining spectral sensitivity to the environment including hydrogen bonds and water molecules and a better understanding of the observed experimental bands.



1. INTRODUCTION

Infrared spectroscopy has been a powerful method for generating fingerprints of molecular structure for a long time. It can be used an as analytical tool to probe functional groups with their specific vibrational frequencies and thus distinguish between isomers. Taking advantage of the sensitivity of these frequencies to their local environment, it can also be used to address the conformations of flexible molecules. Extended motifs are amenable to fingerprinting as well, as extensively shown for proteins.[1] Although the technique has been established for a long time, its development remains very active in the condensed phase, with multidimensional and ultrafast techniques rapidly evolving. It is equally vivid in the gas phase, with both double resonance and photodissociation techniques developed and applied to both neutral and ionic species in the past decade. IR-UV double resonance spectroscopy is carried out on very cold ions leading to high resolution by exploiting resolved UV transitions to monitor infrared spectra for conformers or isomers one at a time. InfraRed Multiple Photon Dissociation (IRMPD) spectroscopy allows for recording, at either low or room temperature, infrared action spectra of gaseous molecular ions that are formed and mass-selected by coupling a laser and a mass spectrometer. It has been applied successfully to organic, inorganic, and biological species, enabling the distinction between isomers and, in some instances, between conformers.[2−4]

The vibrational frequencies of IR modes involving CO, NH, and OH groups are very sensitive to their environment, especially the hydrogen bond network. IR spectroscopy is then able to provide structural signatures of biologically relevant molecules.[5] However, assigning IR bands to specific vibrational modes requires most of the time molecular modeling. This is principally done by quantum mechanics calculations in the static double harmonic approximation. However, in these calculations, temperature and anharmonic effects such as frequency shifts and band broadening as well as band combinations and overtones are not taken into account. Traditionally, scaling factors applied to the computed frequencies are used to reduce the difference with experimental values.[6,7] Furthermore, intense efforts were made to explicitly include anharmonic effects in quantum chemistry vibrational frequency calculations.[8−14] Most of the methods used are based on the vibrational self-consistent field (VSCF) approximation to solve the vibrational Schrödinger equation.[8,9] The potential energy is developed in Taylor series with normal coordinates around the equilibrium geometry. A second-order correction can be added in the perturbation theory framework (VPT2).[10] Vibration configuration interaction (VCI) is a variational way to attain high precision but with a high computational cost drawback.[15] These powerful approaches provide an accuracy around 5 to 15 cm^{-1} but remain limited to small-size systems.[16,17]

An arguably more direct way to include both anharmonicity and temperature effects in the computed spectrum is to perform molecular dynamics simulations. Born−Oppenheimer (BOMD) and Car−Parrinello (CPMD) molecular dynamics associated with the Fourier transform of the Dipole AutoCorrelation Function (DACF) were shown to be an effective approach to compute IR spectra for small molecules of biological interest

Received: February 18, 2014
Published: June 3, 2014

   dx.doi.org/10.1021/ct5004065 | *J. Chem. Theory Comput.* 2014, 10, 3190−3199

both in gas-phase and in solution.[18−23] For example, CPMD simulations for the gas phase protonated dialanine peptide highlighted the dynamical exchange between different isomers via proton transfer at the N-terminus. The flexibility and the dynamics of the peptide were characterized by finite temperature spectra.[24] These calculations give access directly to frequencies, band shapes, and intensities. However, as simulations are time-limited by the size on the system, one can expect that whenever extensive conformational sampling is required to reproduce the experimental conformational fluctuations, CPMD calculations become unaffordable with standard computational resources.[25] DFTB-based dynamics has been used as well.[26] On the other hand, classical molecular dynamics (MD) simulations with force fields have been also used, followed by the DACF treatment, to compute IR spectra.[27−29] However, classical point charge force fields are usually unable to predict frequency shifts due to chemical environment.[30]

Development of modern force fields is a means of correcting this problem through the use of high-level electrostatics and full treatment of polarization effects.[28,31−33] Along such lines, the AMOEBA force field stands as one of the most efficient to reproduce structural, dynamical, and thermodynamical properties of hydrated cations and biomolecules.[34−40] We present in this paper the implementation of the DACF approach associated with the AMOEBA polarizable force field and some applications to biomolecules. The AMOEBA force field has proved to be efficient to handle dynamic simulations and yield structural and energetics data for cations and molecules of biological interest.[34,35,38] We present here a step toward its extension to the simulation of IR spectra. This implementation highlights the necessity to extract new parameters for the AMOEBA force field to be able to simulate accurate IR spectra. Various systems are tested in comparison with both ab initio and experimental spectra, from *N*-methylacetamide (NMA) as a model of the peptide bond to more biologically relevant systems such as helix-shaped peptides. Moreover, one new experimental IRMPD spectrum is provided for the sodiated $(NMA)_2$ dimer and compared with the AMOEBA-DACF simulations.

2. COMPUTATIONAL STRATEGY

2.1. Quantum Chemistry Calculations. Structures, energetics, and IR spectra were computed in the static double harmonic approximation at both ab initio and DFT levels for comparison to AMOEBA-DACF. The levels were chosen to ensure high accuracy. RI-MP2 and RI-CC2 calculations were carried out with the TURBOMOLE program package.[41−43] Geometry optimizations were performed using the resolution-of-the-identity[44,45] (RI) approximation with split valence plus polarization (SVP), double polarized valence-triple-ζ (TZVPP), cc-pVTZ, and aug-cc-pVTZ basis sets. Vibrational frequencies were calculated via normal-mode analysis at the same level as the geometry optimization without the RI approximation. The vibrational frequencies were scaled by a factor depending on the method (see Table 2 for isolated NMA), comparable to the ones used in previous studies.[46] Furthermore, DFT calculations were performed on the $(NMA)_2$ dimer using the B3LYP functional with and without dispersion corrections[47] and different basis sets. The corresponding scale factors are mentioned in Table 3. A single scale factor was used per method for all basis sets for the sake of simplicity. Although this has some influence on absolute frequency values, it has little effect on the N−H stretch frequency shift in $(NMA)_2$ vs NMA, our main focus for this system.

2.2. Dipole Autocorrelation Function Fourier Transform Approach (DACF). An alternative approach toward vibrational spectra is the direct analysis of system dynamics through time series analysis. For a stationary random process, the Wiener-Khinchin theorem[48] relates the power spectrum to the Fourier transform of the autocorrelation function (see the SI for equations). In this framework, by choosing the dipole moment as the observable, the power spectrum obtained from the Fourier transform of its autocorrelation function (named DACF for Dipole Autocorrelation Function) is an infrared vibrational spectrum.[49] For applications such as simulated IR spectra from MD trajectories only a finite set of sampling points is available. The autocorrelation function applied to the dipole moment $\vec{\mu}$ is defined as

$$C(m) = \frac{1}{n}\sum_{i=1}^{n} \vec{\mu}(m+i)\cdot\vec{\mu}(i) \text{ with } m \leq n \tag{1}$$

where m is the time-lag of the autocorrelation function, n is the total number of time steps, and $\vec{\mu}(i)$ is the dipole at time step i. The spectral density function $I(\omega)$ is then obtained by computing the Fourier transform of this discretized autocorrelation function $C(m)$.

The Tinker program[50] was used throughout. It was modified in order to calculate and store the instantaneous dipole moment of the system under consideration resulting in an increase of the global CPU time depending linearly upon the number of atoms. The Fourier program[51] was used to calculate the Fourier transform of the dipole autocorrelation function in order to obtain the power spectrum. Classical time evolution of the nuclei induces an approximation to the calculation of the spectral density. The formula for the coefficient of infrared absorption by unit of length in the approximation of the electric dipole is given as a function of the spectral density by

$$\alpha(\omega) = \frac{4\pi^2\omega}{3V\hbar cn(\omega)}(1 - e^{-\beta\hbar\omega})I(\omega) \tag{2}$$

with V being the volume of the sample, and $n(\omega)$ is the index of refraction at temperature $T = 1/(k_B\beta)$. For gas phase spectroscopy, n is close to 1. In the harmonic approximation for the quantum correction, the final absorption coefficient α_{HA} has the expression[52]

$$\alpha_{HA} = \frac{4\pi^2\omega^2}{3Vcn(\omega)}I_{CL}(\omega) \tag{3}$$

More details about the quantum corrections can be found in refs 53 and 54 and in the SI. A final smoothing of the spectra was achieved through a LOESS procedure.

2.3. The AMOEBA Force Field and Its Parameters. *2.3.1. Description of the Force Field.* The AMOEBA polarizable force field was chosen for its good performances in reproducing cluster, liquid, and solid bulk properties.[35−38,55−57] It is also known to be efficient for computational X-ray crystallography.[58] It has previously been used successfully to determine solvation free energies of cations in liquid phases,[35,59] to calculate protein−ligand binding free energies from explicit and implicit solvent simulations,[60,61] to investigate the conformations and stabilities of model peptides,[40,62,63] and to establish the role of polarization effects on energies in oligoacene crystals.[64] In the context of the present study in the gas phase, AMOEBA also has the appealing feature that it is entirely based on ab initio data, avoiding experimental information that is sometimes not available.

dx.doi.org/10.1021/ct5004065 | J. Chem. Theory Comput. 2014, 10, 3190−3199

Furthermore, due to high-level multipolar electrostatics and explicit treatment of polarization effects, this force field is able to take into account the physical effects involved in charged species. Previous work reported that this type of force fields produces a clear improvement in the electrostatic representation leading to accurate potential energy surfaces with certainly an impact on geometries and IR spectra.[65,66]

The analytical form of the force field is based on second-order exchange-perturbation theory of intermolecular forces. The electrostatic component incorporates permanent charges as well as dipole and quadrupole moments on each atom as derived from quantum mechanical calculations.[59] Many-body polarization effects are explicitly treated using a self-consistent atomic dipole polarization procedure.[62] A polarization-damping scheme is used via a smeared charge distribution as proposed by Tholé.[67] Repulsion-dispersion interactions between pairs of nonbonded atoms are represented by a buffered 14−7 potential.[68] The force field is fully flexible as is mandatory for calculation of vibrational spectra. The intramolecular valence terms consist in bond stretchings, angle bendings, and torsions with the MM3 force field energy form, including anharmonicity effects with higher-order deviations from ideal bond lengths and angles.[69] Additional terms are used such as out-of-plane bendings and valence crossing terms to model the coupling between bonds and angles.

Force field bonded interaction parameters are typically optimized in order to lead to structural, energetic, and vibrational properties as close as possible to available experimental or quantum chemical references.[70−72] As discussed in section 2.1, a scaling of quantum mechanics normal modes is a practical way to reduce the error between computed and experimental vibrational frequencies. Force fields whose harmonic frequencies reproduce anharmonic experimental at nonzero temperature or scaled quantum mechanic spectra implicitly compensate anharmonic effects by lowering the associated force constants. Consequently, anharmonicity brought by the DACF approach would add to this anharmonicity implicitly present in the adapted force constants and would consequently be overestimated. Considering the lightness of hydrogen nuclei, vibration modes involving those elements can be strongly affected by anharmonicity effects, even for strong bond-stretching modes.[16] Following this idea, we found that for vibrational spectra simulated by AMOEBA-DACF, the N−H stretching frequency is improved if the corresponding force constant is decreased. As in the original AMOEBA force field, different atom types and classes were kept for NMA and amino acids; this force constant was decreased from 542 to 500 kcal/mol/Å^2 for the NMA molecule and from 487 to 469 kcal/mol/Å^2 for amino acids.

2.3.2. Determination of New Multipoles for Improvement of IR Spectra. Atomic multipoles play a key role in a vibration mode sensitivity to its environment and for that reason they have been reoptimized for infrared simulation purpose for a set of molecules. Improved atomic multipoles were extracted from ab initio calculations thanks to the distributed multipole analysis (DMA) developed by Stone.[73,74] The atomic local coordinate frame is defined for each site for rotation of the multipole moments.[75] The main idea is to distribute the charge density from the quantum wave function into local multipole moments. Usually in AMOEBA, the DMA procedure is limited to the determination of atomic charges, dipoles, and quadrupoles using the GDMA program.[76] To be consistent with the general set of parameters in the force field, DMA were performed on MP2 wave functions with cc-pVTZ basis sets calculated with the Gaussian program package.[77] A new procedure to extract multipoles for biological molecules was recently proposed with the aim to determine accurate free energies.[39] This procedure includes a fit of the electrostatic potential after the DMA procedure. However, in the case of gas phase spectroscopy, we found that the classic DMA procedure was successful without need for this extra step.

The original Gauss-Hermite algorithm in the GDMA program was used to obtain the multipoles from Gaussian functions. The program applies atomic radii to weight the atomic sites to have a better convergence of the multipolar expansion. Two different radii were used: 0.65 Å for every atom except for hydrogen and a smaller radius of 0.31 Å for hydrogen atoms as found to be necessary to reproduce atomic charges which are chemically intuitive.[76] The resulting multipoles are converted from spherical harmonics to parameters expressed in atomic local Cartesian frames using AMOEBA conventions.[38]

To allow for a better representation of hydrogen bonds and electrostatic interactions in the gas phase, and consequently an accurate reproduction of vibrational frequencies, the multipoles of the isolated molecules were optimized following the above-mentioned procedure, i.e. the NMA molecule and the protonated phosphoserine amino acid, pSerH$^+$. The atomic multipoles of NMA were directly used to verify their transferability in different chemical environments. In the case of peptides, a similar procedure was carried out by dividing the peptide in subsystems consisting in amino acids with either natural or caped terminations. Parameter transferability imposes total charge to be an integer for any residue, except for a terminal one. Another condition was applied on terminal residues to ensure an overall integer charge resulting from the combination of terminal residues. This required small charge adjustments after multipole extraction. For the needs of the cases described herein, three residues were treated: Phe and Ala in central position and Lys in a protonated C-terminal position.

For such new set of multipoles (provided in the Supporting Information), the force field is labeled *AMOEBA_mod* in the following sections. The MD spectra obtained at different temperatures for the phosphoserine amino acid reproduce the main experimental bands with a broadening of the bands by increasing the temperature. To limit the discussion to peptide models here, these results about phosphate modes are described in detail in the SI.

2.4. Simulation Details. For the DACF procedure, molecular dynamics trajectories were propagated with a time step of 0.1 fs, and dipole moment was calculated at each step. Thus, the time interval for dipole collection is well below the upper limit required to compute frequencies in the N−H stretch range. The use of a thermostat to control the temperature may induce a bias on the trajectories. Three thermostats have been tested on the NMA molecule, and both Berendsen and Andersen ones revealed artifacts for our simulations. Accordingly, the temperature was maintained with a Nosé−Hoover[78,79] thermostat for 100 ps after which a trajectory in the microcanonical ensemble was performed for 200 ps. During this period, the system is thermally isolated, and the trajectory is not biased by any thermostat, except potentially for the starting point. This starting point is defined by both atomic positions and velocities. Finally, DACF spectra have been averaged over 10 trajectories to limit artifacts caused by the starting conformation. The Tinker package was used for all simulations.[50]

The complex processes involved in action spectroscopy make it fundamentally different from linear absorption spectroscopy. This has been discussed in detail in ref 80 for IRMPD. It turns out

that absorption frequencies are generally very similar.[81] On the other hand, absolute intensities can be significantly different. Therefore, comparison of experimental efficiencies of action spectra to calculated absorption intensities and band widths may only be qualitative for some systems. This may be improved for example by taking explicitly into account the multiple photon and slow heating character of IRMPD,[82] however at much increased computational burden.

3. RESULTS AND DISCUSSION

3.1. *N*-Methylacetamide and Derivatives. *3.1.1. Hydrogen Bond Directionality and Interaction Energies.* AMOEBA multipoles were extracted from ab initio calculations through the procedure mentioned in section 2.3.2. The transferability of the parameters was tested on three complexes including NMA-Na^+ for an interaction with a cation, NMA-H_2O and $(NMA)_2$ with one hydrogen bond between the N−H of the amide group and the oxygen of the water molecule or between one N−H and one C═O, respectively (Figure 1). The latter case is especially interesting because the same multipoles are used for both NMA molecules even if they act as hydrogen bond donor or acceptor. Interaction energies for these complexes are presented in Table 1, calculated as the energy difference between the complex and its fragments (taken at the geometry of the complex). For all three systems, accurate values of the interaction energy are obtained in comparison with MP2 values, with some significant improvement compared to standard AMOEBA especially for the Na^+ interaction. New NMA multipoles yield also to a better directionality of the hydrogen bonds for $(NMA)_2$ and NMA-H_2O with a good agreement for hydrogen bond distances and angles. Electrostatic interactions are better reproduced for NMA-Na^+ with a CO···Na^+ distance of 2.02 Å at the MP2/aug-cc-pVTZ level, 2.09 Å with AMOEBA and 2.01 Å with *AMOEBA_mod*. In general, very good agreement is obtained for both geometries and interaction energies using the new multipoles.

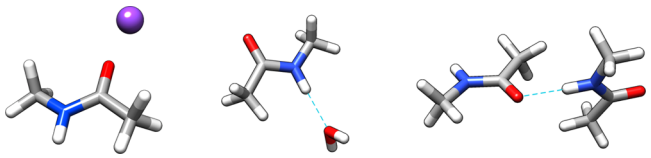


Figure 1. MP2/cc-pVTZ optimized geometries for NMA-Na^+, NMA-H_2O, and $(NMA)_2$.

Table 1. Interaction Energy of the NMA Complexes Presented in Figure 1 (kcal·mol^{-1})

system	MP2/aug-cc-pVTZ	MP2/cc-pVTZ	AMOEBA	AMOEBA_mod
NMA-Na^+	−44.1	−42.2	−37.7	−45.6
NMA-H_2O	−6.7	−7.3	−5.2	−5.9
$(NMA)_2$	−11.7	−10.4	−9.4	−12.4

3.1.2. Vibrational Frequencies for NMA. In a peptide, the structure is mainly defined by the backbone geometry. Consequently, the amide modes of peptide bonds provide a strong insight into the tridimensional structure and especially for the determination of secondary structures. Among these amide modes, two frequency ranges are to be distinguished, i.e. N−H stretches (amide A) usually in the 3200−3600 cm^{-1} region, and in the 1200−1800 cm^{-1} range, the amide I, II, and III bands. At lower frequencies, other amide bands may be defined but will not be discussed in this paper. The IR spectrum of isolated NMA has been obtained previously under different experimental conditions, including argon matrix at 20 K,[83] molecular jet,[84] and gas phase near room temperature.[85] The data which are most relevant to the present work are provided in Table 2. As NMA is a simple model of a peptidic bond, high level static quantum chemistry calculations can be performed. Furthermore, its IR spectrum was intensively scrutinized using different MD approaches, providing an appropriate testing ground. The comparison with various sets of results is described below. Only the trans NMA conformation was studied even if in some experiments cis and trans conformers were observed.

Static Harmonic Spectrum. Experimental frequencies for amide modes in NMA are compared with values obtained by normal coordinate analysis using MP2 and AMOEBA (Table 2). The CC2 level using the same basis sets was checked to lead to similar results than MP2 as it will be discussed later for the NMA dimer. Both MP2 and AMOEBA reported frequencies are harmonic, and the values are provided with no other correction than the use of a scale factor for MP2 frequencies. Amide frequencies at the MP2 level are sensitive to the basis set but globally fall in the experimental range. It is interesting to note that MP2/SVP is a good compromise between accuracy and computational time. For amide I and II modes, the inclusion of the new multipoles in AMOEBA leads to accurate values in comparison with ab initio or experimental data. These results indicate that AMOEBA is reliable for simulating low temperature IR spectra especially below 2000 cm^{-1}. As mentioned by Keiderling and Kubelka,[85] combinations with methyl vibrations make the amide III band difficult to predict. It will not be discussed herein. Regarding the amide A mode, the frequency is overestimated by 45 cm^{-1} which may be attributed to strong anharmonicity. Comparison with experiments thus requires the simulation of spectra including temperature effects, as described in the next section.

Temperature Dependent Spectrum. The DACF procedure was applied at different temperatures from 50 to 413 K. The length of the simulation of 200 ps allows taking into account the possible effects of a dynamical conformational sampling in the spectrum. The resulting MD spectra for three temperatures in

Table 2. Isolated NMA Frequencies (cm^{-1})^f

	exp. refs 84 and 85	CPMD^a ref 18	MP2^b SVP	MP2^c cc-pVTZ	MP2^d aug- cc-pVTZ	AMOEBA original	AMOEBA_mod new multipoles^e
amide A	3505	3344	3520	3500	3514	3488	3550
amide I	1713−1731	1609	1724	1705	1700	1703	1716
amide II	1497−1500	1458	1494	1511	1520	1541	1523

^{a}MD simulations at 20 K using the BLYP density functional. ^{b}MP2 frequencies scaled by 0.943. ^{c}MP2 frequencies scaled by 0.94 (>3000 cm^{-1}) and 0.96 (<2000 cm^{-1}). ^{d}MP2 frequencies scaled by 0.95 (>3000 cm^{-1}) and 0.97 (<2000 cm^{-1}). ^{e}Multipoles from MP2/cc-pVTZ wave function. ^{f}AMOEBA and AMOEBA_mod frequencies correspond to static harmonic values.

dx.doi.org/10.1021/ct5004065 | J. Chem. Theory Comput. 2014, 10, 3190−3199

comparison with experiments are provided in Figures 2 and 3 in the 1400−1800 and 3300−3600 cm^{-1} ranges, respectively. In the

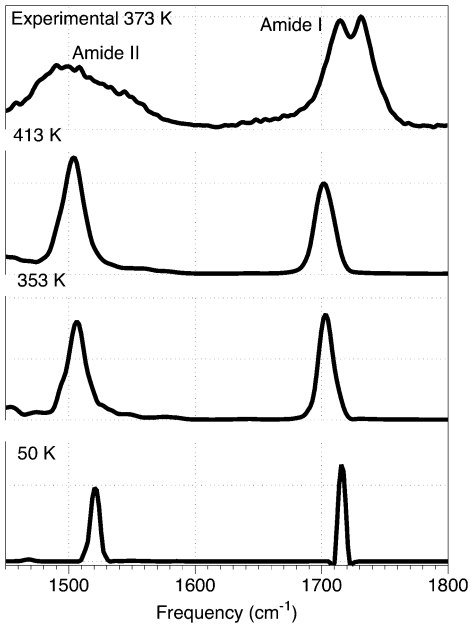


Figure 2. Isolated NMA spectrum in the 1400−1800 cm^{-1} range using the polarizable DACF procedure at 50 K, 353 K, and 413 K. The experimental spectrum at ca. 373 K (top) is from ref 85.

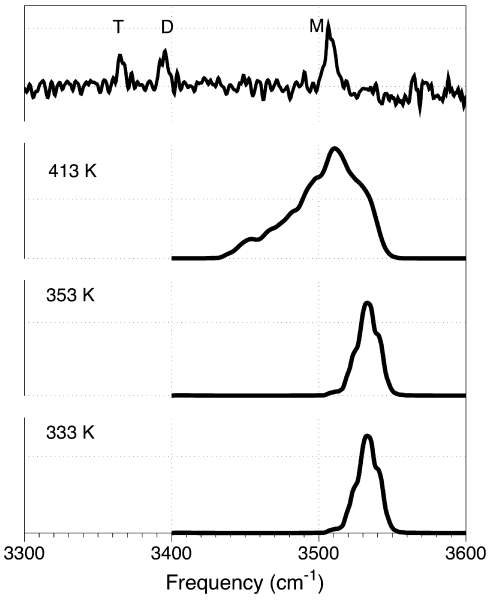


Figure 3. Isolated NMA spectrum in the 3300−3600 cm^{-1} range using the polarizable DACF procedure at 333 K, 353 K, and 413 K. The experimental spectrum at ca. 333 K (top) is from ref 84. M stands for NMA monomer, D for dimer, and T for trimer.

experimental spectrum, the amide I is a doublet and the amide II is broad, both probably due to the presence of cis and trans conformers of NMA. With the aim to use NMA as a model for the peptide bond, only the trans conformer was used in the simulations leading to thinner bands than in experiments. In the amide I and II range, the major effect of temperature is the broadening of the bands; shifts of 15−20 cm^{-1} are discernible on the band positions. Band maxima at 50 K are close to the static frequencies described previously. The broadening and red shift of the amide II band bring it in significantly better agreement with experiment at 353 or 413 K than at 50 K, although its width remains smaller than experiment. Broadening is smaller for the amide I band in agreement with experiment; however, in this case the red shift leads to a slightly worse agreement for the frequency maximum. Overall, the spectrum computed at 413 K reproduces rather well the experimental spectrum recorded at an approximated temperature of 373 K.[85] In the 3300−3600 cm^{-1} region, only the trans conformation is observed experimentally. The position of the N−H stretch band maximum as well as its broadening appear to be much more sensitive to temperature. The experimental band maximum is used here as a reference for the calibration of the force field to account for temperature and anharmonicity. Thus, comparison of computed and experimental spectra is deferred to larger systems (see below).

These results can be compared with available results from MD simulations in the literature by DFT-based Car−Parrinello MD using the BLYP functional,[18] semiempirical Born−Oppenheimer MD (SEBOMD),[20] or classical polarizable MD.[29] The SEBOMD approach with the PM3 Hamiltonian offers the possibility to run long-time scale simulations. The results lead to a considerable overestimation of the amide I frequency and an underestimation of the amide II and amide A frequencies for gas phase NMA. However, the goal of this study was to predict solvent effects on the vibrational frequency of amide bonds, and a general good agreement with experiments was found on frequency shifts. A comparison with the AMBER03 force field highlights that the latter is not able to reproduce the correct ordering of amide I and II frequencies.[29] The authors support the need for a correct description of electronic properties such as electrostatic and polarization interactions for an accurate modeling of IR spectra. Quasi harmonic frequencies computed with CPMD and the BLYP functional are found to be not very accurate, with deviations ranging from −50 cm^{-1} for the amide II frequency to −160 cm^{-1} for the amide A frequency, before using a scaling factor to compensate for BLYP intrinsic errors.[18] As for SEBOMD, their final results in solution give a satisfactory agreement with experiments. Finally, Tavan et al. reported the MD NMA spectrum in gas phase and in solution using a point charge polarizable force field with parameters extracted from DFT/BP86 data.[28] This spectrum requires the use of a scaling factor to be brought into good agreement with the gas phase experiments.

3.1.3. Vibrational Frequencies for $(NMA)_2$ and NMA Complexes with Na^+. Hydrogen-Bound Dimer of NMA. With the aim to estimate the effects of intermolecular interactions on the vibrational frequencies, the isolated NMA dimer has been studied. The importance of atomic polarizability and the truncation of atomic multipoles on the length of the hydrogen bond and the corresponding N−H stretch red shift is shown in Table S1 using static and harmonic approximations. The absence of anharmonic terms in the potential restrains bond stretching energy profiles, which has a crucial impact on the effects produced by the environment on the static harmonic vibrational spectrum. The obtained frequency shift is very small in comparison to the experimental one (<10 cm^{-1}). Nonbonded interactions modify the Hessian, even for an isolated molecule of NMA. Therefore, an electrostatic truncation to atomic point dipoles or charges results in changes in vibrational frequencies at 0 K. The removal of atomic quadrupoles leads to a stronger hydrogen bond that is shortened from 1.89 to 1.86 Å, as also reflected by the red shift. An additional removal of atomic dipoles

dx.doi.org/10.1021/ct5004065 | J. Chem. Theory Comput. 2014, 10, 3190−3199

has however an antagonist effect, as the corresponding hydrogen bond is importantly extended to 2.04 Å. The directionality of the hydrogen bond is also strongly affected by the truncation of the multipoles. Whether or not it is associated with a truncation of the permanent electrostatic description, the removal of polarization causes a lengthening of the hydrogen bond as well as a halving of the separation between the free and the bonded amide A frequencies.

The spectrum was then computed at finite temperature and compared with experiments (see Figure 4). The dimer band at

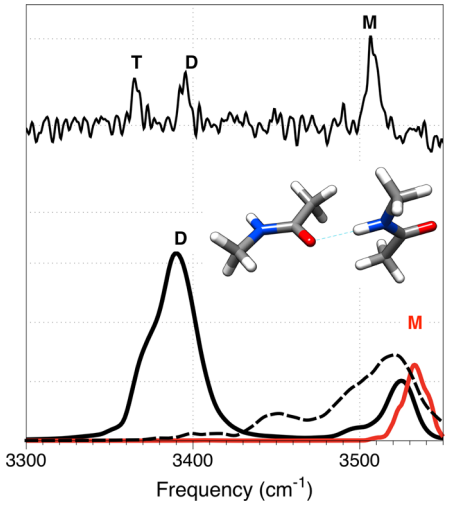


Figure 4. Comparison between the NMA dimer (D, black line) and monomer (M, red line) MD spectra computed at 353 K in the 3300−3600 cm^{-1} range and the experimental spectrum at ca. 333 K (top) from ref 84. The dashed-line corresponds to the spectrum obtained for the NMA dimer from simulations without polarization (see text).

3385 cm^{-1} is weak because of the low abundance of the dimer relative to the monomer in experiment. IR spectra for the NMA dimer were obtained in N_2 and Ar matrices at low temperature[86] and using jet FTIR in the gas phase at ca. 200 K;[84] however, in the latter case it was assumed that the temperature may vary according to the mode. In the NMA dimer, the positions of the free N−H and C═O stretching modes are very close to the values in the monomer. A red shift is obtained experimentally for the amide A and amide I modes corresponding to the N−H and C═O groups involved in the hydrogen bond and can be compared to calculated shifts. The amide I red shift is 32 cm^{-1} in the jet FTIR experiment, and it is reasonably well reproduced at the *AMOEBA_mod* level with the value of 22 cm^{-1}. In the same experiment, the N−H stretch red shift is about 115 cm^{-1}. The *AMOEBA_mod* value from MD simulation of 135 cm^{-1} is only slightly larger, as shown in Figure 4. If polarization is turned off during MD simulations, the red shift is reduced to 72 cm^{-1}. This value of 135 cm^{-1} may be considered as good agreement, all the more as this value is difficult to reproduce accurately by harmonic quantum chemistry calculations even with high level methods. Table 3 presents the results for various levels of calculations using different basis sets. The most appropriate method is DFT/B3LYP using a triple-ζ quality basis set, certainly due to favorable error compensation, in comparison with DFT/B3LYP-D3 or post-Hartree−Fock methods. These results highlight the importance of taking into account anharmonicity and polarization in the calculation of N−H stretching modes.

Table 3. Computed Harmonic Frequencies (cm^{-1}) of $(NMA)_2$ Using Various Levels of Quantum Chemistry[d]

method	basis set	free N−H	bound N−H	$\Delta\nu$
MP2[a]	SVP	3497	3352	145
	TZVPP	3492	3316	176
	cc-pVTZ	3503	3328	175
	aug-cc-pVTZ	3480	3315	165
CC2[b]	SVP	3501	3342	159
	TZVPP	3480	3297	183
	cc-pVTZ	3496	3318	178
	aug-cc-pVTZ	3468	3298	169
B3LYP-D3[c]	SVP	3488	3326	161
	TZVPP	3498	3349	149
	cc-pVTZ	3495	3345	151
	aug-cc-pVTZ	3594	3344	151
B3LYP[c]	SVP	3488	3336	152
	TZVPP	3498	3377	121
	cc-pVTZ	3495	3365	130
	aug-cc-pVTZ	3496	3379	117
AMOEBA_mod		3525	3390	135
exp.[84]		3510	3395	115

[a]Scale factor of 0.943. [b]Scale factor of 0.95. [c]Scale factor of 0.96. [d]AMOEBA_mod values from MD simulations (Figure 4).

Complexes of NMA with the Sodium Cation. An experimental IRMPD spectrum was recorded for the sodiated NMA dimer $(NMA)_2$-Na^+ in the CLIO setup at room temperature in the 1200−1800 cm^{-1} range, in order to identify the frequency shifts resulting from the complexation to a sodium cation.[87] Both experimental and DACF spectra at 50 and 300 K are displayed in Figure 5. Amide I and amide II bands are easy to identify near 1650 and 1550 cm^{-1}, respectively. The direct complexation of the two C═O groups to the sodium cation leads to a red shift for the C═O stretching bands of 62 cm^{-1} experimentally. This value is well reproduced by the force field with a shift of 53 cm^{-1}. If polarization is switched off in the model,

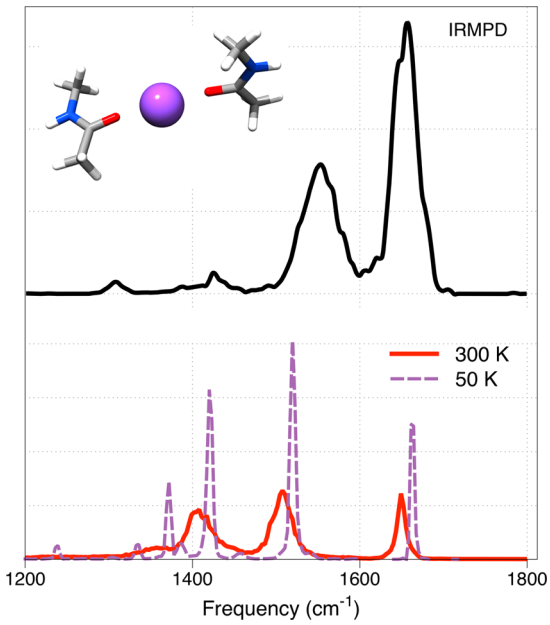


Figure 5. Computed MD spectrum of the $(NMA)_2$-Na^+ complex at 50 and 300 K in comparison with the spectrum obtained from IRMPD experiments in the 1200−1800 cm^{-1} range.

dx.doi.org/10.1021/ct5004065 | J. Chem. Theory Comput. 2014, 10, 3190−3199

the red-shift is limited to 40 cm^{-1}. However, the experimental blue-shift of the amide II band in comparison with bare NMA is not reproduced by this version of the force field. This will deserve more calibration in the future. MD spectra have been obtained at different temperatures to point out the effect on band positions and intensities. At low temperature, the red-shift of amide I is only around 40 cm^{-1}. Interestingly, the relative intensities of amide I and II bands are reverse in comparison with experiments. The band at ca. 1420 cm^{-1} is also very intense. The intensities are in better agreement with experiments at 300 K with broadening of the bands especially in the 1320−1450 cm^{-1} range. Simulations at temperatures higher than 300 K have shown a larger broadening in this region.

Previous gas-phase IR spectroscopy studies in the 3000−3800 cm^{-1} range investigated the effect of strong electrostatic interactions in $M^+(NMA)(H_2O)_n$ ($M^+ = K^+$ or Na^+, n = 0−3).[88] The experimental temperature was estimated to be ca. 250−300 K. MD spectra were obtained from AMOEBA simulations for these complexes and are presented in Figure 6 (they can be compared to the experimental ones in Figure 2 of ref 88). For n = 0, the cation interacts with NMA via the carbonyl group that shifts the position of the N−H stretch band at 3483 cm^{-1} experimentally. This shift is well reproduced by AMOEBA leading to a band position at 3485 cm^{-1}. For n = 1 and n = 2, three bands are present, i.e. one for N−H stretch mode as for $Na^+(NMA)$, two for the symmetric and antisymmetric O−H stretches of water. The spectra can be attributed to a structure in which NMA and the water molecules interact individually with the sodium ion.[88] This results again in shifts for the N−H and O−H symmetric and asymmetric stretch bands that are reproduced by the simulations. For n = 3, a competition between hydrogen bonding and direct electrostatic interaction with the cation begins to appear. In the case of the sodium complex, no signature of water−water hydrogen-bonded isomers is present, leading to a complex with NMA and the water molecules in the first coordination shell of Na^+. The broadening of the O−H stretch bands arises from the dynamics of the water molecules around the Na^+. The resulting N−H and O−H signatures from n = 0 to n = 3 are in agreement with the calculated spectrum.

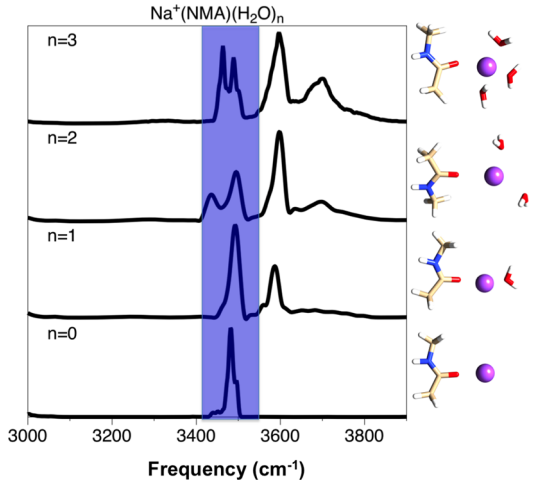


Figure 6. Computed MD spectra of the $Na^+(NMA)(H_2O)_n$ (n = 0, 1, 2, 3) complexes at 250 K in the 3000−3800 cm^{-1} range.

3.2. Conformations of the Helix Model Ac-Phe-$(Ala)_{10}$-Lys-H^+. Isolated polyalanine-based peptides have been investigated in the gas phase during the past decade.[21,25,89−93] Ala has the highest helix propensity of all amino acid residues in solution,[94] and stable polyalanine helices have also been detected in the gas phase.[89−91,93] One of the aims of those studies is to improve the knowledge of the intrinsic helical trend of polyalanine peptides. Their hydrogen-bond network generates characteristic features in infrared spectra which are therefore a useful tool for revealing structural differences. Modeling the structure and dynamics of such peptides requires accurate description of hydrogen bonds and the relative stability of the various conformations. Dynamical approaches have been carried out recently to introduce both dynamics and temperature effects in the simulations.[25,95]

Rizzo et al. published low temperature, highly resolved spectra for Ac-Phe-$(Ala)_5$-Lys-H^+ and Ac-Phe-$(Ala)_{10}$-Lys-H^+[90,91] in which Phe is used as a chromophore for IR-UV double resonance spectroscopy. The protonated Lys at the C terminus stabilizes the helix macrodipole and coordinates to three carbonyl groups, as a key element favoring helical structures. The helical structure of Ac-Phe-$(Ala)_n$-Lys-H^+ was found to become favored relative to globules with growing n.[21,91,96] Rizzo et al. were able to detect two stable conformers for Ac-Phe-$(Ala)_{10}$-Lys-H^+ (mentioned as conformers c and d in ref 90). They are expected to correspond to helical or partly helical structures, in which peptidic H−N groups bound through C_{10} or C_{13} interactions play a crucial role for helix stabilization. Helical N−H stretches were assigned to the 3320−3350 cm^{-1} range. The main difference between c and d comes from a band at ca. 3380 cm^{-1} in conformer d which is not present in conformer c. The authors invoked possible C_5 or N−H$\cdots\pi$ interaction in order to explain this difference; however, no definitive conclusion could be reached.

Due to the low-temperature conditions of the experiment, this system is not a test for taking into account temperature effects. It is rather a good test for demonstrating intrinsic AMOEBA capabilities to reproduce frequency shifts in peptidic amide groups due to a variety of interactions. MD simulations of ca. 10 ns at 298 K were run in order to explore the conformational landscape, starting from an α-helix conformation of the peptide. A number of low-energy conformations were collected through the trajectory and were reoptimized with the force field. All the resulting minima had a significant helical character. The Lys side chain acts as a global stabilizator of the helix via charge-dipole interactions; however, it is also a local perturbator of the helix as it is a triple hydrogen bond donor. A compromise between these competing effects leads to the Lys side chain being always linked to the carbonyl of the last Ala residue. All structures obtained are found to contain 5 to 8 hydrogen bonds typical of an α-helix, with some having locally either $3_{10}10$- or π-helical character, which is not expected to have a dramatic effect on the IR spectrum. This is consistent with the hypothesis of a largely α-helical structure for this peptide, made by Rizzo et al.[91]

From this group of stable structures, the most populated during the 10 ns was selected, named H_2 below. It is the second lowest in energy, 0.8 kcal·mol^{-1} higher than the lowest found. The statistical occurrence of a single energy basin during the dynamics appears to be more relevant than the energy difference of the individual minima. A second structure was selected on the basis of structural diversity. Although three other structures were found to be essentially degenerate with H_2, their slight differences in types of helices are not expected to be responsible for the formation of the new bands in the second experimental IR spectrum. Thus, we rather picked up the lowest energy structure in which the Lys side chain is oriented differently, forming a C_7 motif, as a reasonable candidate to explain the distinct

dx.doi.org/10.1021/ct5004065 | J. Chem. Theory Comput. 2014, 10, 3190−3199

experimental spectrum. This second structure, named H_1 below, lies 2.4 kcal·mol^{-1} higher in energy than H_2. The two structures are shown in Figure 7. The C_7 interaction exists in H_1 (blue circle in Figure 7) with a hydrogen bond length of 2.0 Å but not in H_2 (red circle).

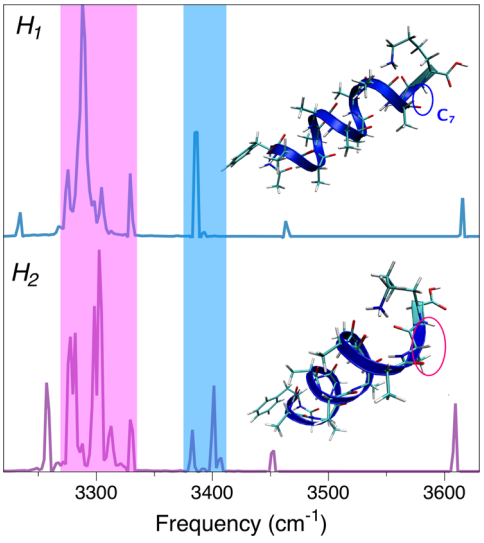


Figure 7. Conformers H_1 and H_1 of Ac-Phe-$(Ala)_{10}$-Lys-H^+ and MD spectra at 10 K. Loss of one C_7-motif at the C-terminus for H_2.

DACF AMOEBA spectra were then computed at 10 K from MD trajectories using these conformations as starting points, in the conditions described in the methodological section. The computed spectra are also shown in Figure 7. Given the very high sensitivity of N−H stretching frequencies to the fine details of the environment, the overall agreement with experimental spectra[91] is very encouraging. The band at ca. 3610 cm^{-1} corresponds to the C-terminal carboxylic acid O−H stretching mode. Its frequency is overestimated by 40 cm^{-1} in comparison with experiment. No force constant calibration was made for this group. We find only one free N−H bond, for Ala_1 (numbering of Ala residues is from N- to C-terminus) with a stretching frequency of 3450 cm^{-1}, in very good agreement with experiment. Computed bands near 3400 cm^{-1}, also very close to experimental ones, were tentatively assigned to those N−H's at the N- and C-termini.[91] We find that the Ala_2 N−H is involved in a C_{10} interaction and that this band lies near 3400 cm^{-1} rather than 3300 cm^{-1} as previously suggested. The Lys peptidic N−H also falls in this range when involved in a C_5 interaction (structure H_2), generating a characteristic band in this non-congested area (see below). The helix massif is found to appear in the 3270−3335 cm^{-1} range. This is ca. 40−50 cm^{-1} red-shifted compared to experiment. It is possible that cooperative effects as well as electrostatic effects are overestimated by the AMOEBA force field. This will deserve further investigation. As in experimental spectra, at least one band stands alone to the blue of the main helix massif, and this is attributed to the Ala_4 N−H. Although the latter is involved in the α-helix, its hydrogen bond is ca. 0.04 Å longer than central C_{13} bonds, due to helix boundary effects. To the red of the helix massif (at ca. 3280 cm^{-1}), we find that two ammonium N−H bonds of the Lys side chain are weakly enough interacting with nearby C═Os that the associated red shift is not as large as for the third. These could account for the experimental band near 3290 cm^{-1}. Also lying in this range is the Lys peptidic N−H when involved in a C_7 motif as found in structure H_1, at variance with H_2 where it forms a C_5 with an associated blue shift to 3380 cm^{-1}. Thus, this C_5/C_7 alternative for the Lys N−H offers a possible explanation for the small band found at 3380 cm^{-1} in one of the experimental spectra but not in the other.

The present work may be regarded as a proof of principle, that interpretation of IR spectra of complex and flexible molecules may be obtained with the combination of DACF and the AMOEBA force field. Our results are consistent with a strong helical character of Ac-Phe-$(Ala)_{10}$-Lys-H^+, as suggested by Rizzo et al.[91] However, a true structural identification, including an explanation for the similar experimental spectra recorded for a polyalanine with the Lys at the N- instead of C-terminus, would require a more exhaustive work that is beyond the scope of this work. We emphasize here that the associated computational burden is light enough that many conformers can be studied for even larger cases.

4. CONCLUSIONS

The present work is the first step in the extension of the polarizable AMOEBA force field to infrared spectroscopy. The original force field was adapted mainly by extracting new multipoles using the DMA procedure. The study is devoted to the environment of the C═O and N−H groups of peptide models. Infrared spectra can be calculated with good accuracy from finite temperature molecular dynamics simulations followed by the Fourier transform of the dipole autocorrelation function. The AMOEBA-DACF approach is able to reproduce a large range of frequencies and a variety of modes such as N−H and O−H stretches, amide I and II bands in peptides, stretching and bending phosphate modes (see the SI). The method allows to distinguish between conformers and to reflect the role of the environment, be it solvent molecules, cations, or hydrogen bonds. The study was focused on the role of the nonbonded terms, specifically atomic multipoles and atomic polarizability, to reproduce experimental IR spectra. The next step would require improvement of other terms when necessary, such as intramolecular parameters from the MM3 formalism. In addition, further work will be devoted to the treatment of strong hydrogen bonds that implies partial covalency and large red shifts.

Theoretical approaches to molecular vibrational spectroscopy are already many. Quantum chemical methods in the static approach, with the double harmonic approximation, is extensively used for molecules with several tens or hundreds of atoms, depending on the ab initio or DFT level chosen. However, as recalled above in the NMA case, the harmonic approximation may lead to severe difficulties even for simple molecules. Extending beyond harmonic approximations may lead to very high accuracy, however only for molecules with up to ca. 10−20 atoms. Classical mechanics methods can hardly compete with the best quantum mechanical methods for high accuracy in the static approach. They are however applicable to molecules of virtually any size, and easily amenable to molecular dynamics simulations, thus to the DACF approach. There is ample literature on the application of such methods to proteins, with some significant success. In particular, long simulations allow for significant sampling and statistical averaging, leading to meaningful comparison to experiments. However, when strong electrostatic or polarization effects are present, simple force fields tend to fail for either absolute frequencies, frequency shifts, or both. Thus, coupling a polarizable atomic multipole-based force field with the DACF approach is particularly appealing. Compared to similar

dx.doi.org/10.1021/ct5004065 | J. Chem. Theory Comput. 2014, 10, 3190−3199

approaches based on DFT, the long simulation capability is expected to be a special strength for highly flexible systems.

Work is in progress to simulate microhydrated cationic and anionic species in order to assess the role of the cluster size on the infrared spectra. The high sensitivity of the shape and position of the bands may serve as a tool to probe the subtle changes in the network of hydrogen bonds involving water molecules.

■ ASSOCIATED CONTENT

Supporting Information

Methodological details about the DACF procedure, additional results about the role of atomic polarizability, atomic multipoles and intramolecular parameters for $(NMA)_2$, results for the protonated phosphoserine and new force field parameters. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: gilles.ohanessian@polytechnique.edu.
*E-mail: carine.clavaguera@polytechnique.edu.

Present Address

[¶]Department of Chemistry, Chemical Theory Center, and Supercomputing Institute, University of Minnesota, 207 Pleasant St. SE, Minneapolis, Minnesota 55455, United States.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

Prof. Martin Suhm and Jan Kubelka are gratefully acknowledged for providing the experimental spectra for NMA. Use of the FT-ICR mass spectrometer at Orsay was supported by the TGE "Spectrométrie de masse FT-ICR à très haut champ" funded by the CNRS Institute of Chemistry. We are grateful to Edith Nicol for carrying out the IRMPD experiments on $(NMA)_2Na^+$ and to the CLIO team for support. The CLIO center received funding from the European Union's Seventh Framework Programme (FP7/2012-2015) under grant agreement no. 312284. This work was granted access to the HPC resources of [CCRT/CINES/IDRIS] under the allocation c2013077128 made by GENCI (Grand Equipement National de Calcul Intensif).

■ REFERENCES

(1) Siebert, F.; Hildebrandt, P. *Vibrational Spectroscopy in Life Science*; Wiley-VCH Verlag GmbH & Co.: KGaA, 2007.
(2) Fung, Y. M. E.; Besson, T.; Lemaire, J.; Maître, P.; Zubarev, R. A. *Angew. Chem., Int. Ed.* **2009**, *48*, 8340−8342.
(3) Kupser, P.; Pagel, K.; Oomens, J.; Polfer, N.; Koksch, B.; Meijer, G.; Helden, G. v. *J. Am. Chem. Soc.* **2010**, *132*, 2085−2093.
(4) Balaj, O.; Semrouni, D.; Steinmetz, V.; Nicol, E.; Clavaguéra, C.; Ohanessian, G. *Chem.—Eur. J.* **2012**, *18*, 4583−4592.
(5) Garand, E.; Kamrath, M. Z.; Jordan, P. A.; Wolk, A. B.; Leavitt, C. M.; McCoy, A. B.; Miller, S. J.; Johnson, M. A. *Science* **2012**, *335*, 694−698.
(6) Merrick, J. P.; Moran, D.; Radom, L. *J. Phys. Chem. A* **2007**, *111*, 11683−11700.
(7) Alecu, I. M.; Zheng, J.; Zhao, Y.; Truhlar, D. G. *J. Chem. Theory Comput.* **2010**, *6*, 2872−2887.
(8) Bowman, J. M. *J. Chem. Phys.* **1978**, *68*, 608−610.
(9) Gerber, R. B.; Ratner, M. A. *Chem. Phys. Lett.* **1979**, *68*, 195−198.
(10) Barone, V. *J. Chem. Phys.* **2005**, *122*, 014108.
(11) Carbonnière, P.; Lucca, T.; Pouchan, C.; Rega, N.; Barone, V. *J. Comput. Chem.* **2005**, *26*, 384−388.
(12) Gohaud, N.; Begué, D.; Darrigan, C.; Pouchan, C. *J. Comput. Chem.* **2005**, *26*, 743−754.
(13) Benoit, D. M. *Front. Biosci.* **2009**, *14*, 4229−4241.
(14) Roy, T. K.; Gerber, R. B. *Phys. Chem. Chem. Phys.* **2013**, *15*, 9468−9492.
(15) Bowman, J. M.; Christoffel, K.; Tobin, F. *J. Phys. Chem.* **1979**, *83*, 905−912.
(16) Carbonnière, P.; Dargelos, A.; Ciofini, I.; Adamo, C.; Pouchan, C. *Phys. Chem. Chem. Phys.* **2009**, *11*, 4375−4384.
(17) Thicoïpe, S.; Carbonniere, P.; Pouchan, C. *Chem. Phys. Lett.* **2014**, *591*, 243−247.
(18) Gaigeot, M.-P.; Vuilleumier, R.; Sprik, M.; Borgis, D. *J. Chem. Theory Comput.* **2005**, *1*, 772−789.
(19) Gaigeot, M.-P. *Phys. Chem. Chem. Phys.* **2010**, *12*, 3336−3359.
(20) Ingrosso, F.; Monard, G.; Hamdi Farag, M.; Bastida, A.; Ruiz-López, M. F. *J. Chem. Theory Comput.* **2011**, *7*, 1840−1849.
(21) Rossi, M.; Blum, V.; Kupser, P.; von Helden, G.; Bierau, F.; Pagel, K.; Meijer, G.; Scheffler, M. *J. Phys. Chem. Lett.* **2010**, *1*, 3465−3470.
(22) Carbonnière, P.; Thicoipe, S.; Very, T.; Assfeld, X. *Int. J. Quantum Chem.* **2012**, *112*, 2221−2230.
(23) Thomas, M.; Brehm, M.; Fligg, R.; Vohringer, P.; Kirchner, B. *Phys. Chem. Chem. Phys.* **2013**, *15*, 6608−6622.
(24) Marinica, D. C.; Grégoire, G.; Desfrançois, C.; Schermann, J. P.; Borgis, D.; Gaigeot, M.-P. *J. Phys. Chem. A* **2006**, *110*, 8802−8810.
(25) Tkatchenko, A.; Rossi, M.; Blum, V.; Ireta, J.; Scheffler, M. *Phys. Rev. Lett.* **2011**, *106*, 118102.
(26) Goyal, P.; Ghosh, N.; Phatak, P.; Clemens, M.; Gaus, M.; Elstner, M.; Cui, Q. *J. Am. Chem. Soc.* **2011**, *133*, 14981−14997.
(27) Fanourgakis, G. S.; Xantheas, S. S. *J. Chem. Phys.* **2008**, *128*, 074506.
(28) Schultheis, V.; Reichold, R.; Schropp, B.; Tavan, P. *J. Phys. Chem. B* **2008**, *112*, 12217−12230.
(29) Schropp, B.; Wichmann, C.; Tavan, P. *J. Phys. Chem. B* **2010**, *114*, 6740−6750.
(30) Cézard, C.; Rice, C. A.; Suhm, M. A. *J. Phys. Chem. A* **2006**, *110*, 9839−9848.
(31) Kumar, R.; Wang, F.-F.; Jenness, G. R.; Jordan, K. D. *J. Chem. Phys.* **2010**, *132*, 014309.
(32) Liu, J.; Miller, W. H.; Fanourgakis, G. S.; Xantheas, S. S.; Imoto, S.; Saito, S. *J. Chem. Phys.* **2011**, *135*, 244503.
(33) Voora, V. K.; Ding, J.; Sommerfeld, T.; Jordan, K. D. *J. Phys. Chem. B* **2012**, *117*, 4365−4370.
(34) Ren, P.; Ponder, J. W. *J. Phys. Chem. B* **2003**, *107*, 5933−5947.
(35) Grossfield, A.; Ren, P.; Ponder, J. W. *J. Am. Chem. Soc.* **2003**, *125*, 15671−15682.
(36) Piquemal, J.-P.; Perera, L.; Cisneros, G. A.; Ren, P.; Pedersen, L. G.; Darden, T. A. *J. Chem. Phys.* **2006**, *125*, 054511.
(37) Wu, J. C.; Piquemal, J.-P.; Chaudret, R.; Reinhardt, P.; Ren, P. *J. Chem. Theory Comput.* **2010**, *6*, 2059−2070.
(38) Ponder, J. W.; Wu, C.; Ren, P.; Pande, V. S.; Chodera, J. D.; Schnieders, M. J.; Haque, I.; Mobley, D. L.; Lambrecht, D. S.; DiStasio, R. A.; Head-Gordon, M.; Clark, G. N. I.; Johnson, M. E.; Head-Gordon, T. *J. Phys. Chem. B* **2010**, *114*, 2549−2564.
(39) Ren, P.; Wu, C.; Ponder, J. W. *J. Chem. Theory Comput.* **2011**, *7*, 3143−3161.
(40) Semrouni, D.; Ohanessian, G.; Clavaguéra, C. *Phys. Chem. Chem. Phys.* **2010**, *12*, 3450−3462.
(41) TURBOMOLE V6.1 2009, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989−2007, TURBOMOLE GmbH, since 2007. Available from http://www.turbomole.com (accessed December 2009).
(42) Ahlrichs, R.; Bär, M.; Häser, M.; Horn, H.; Kölmel, C. *Chem. Phys. Lett.* **1989**, *162*, 165−169. For the current version, see: http://www.turbomole.com (accessed December 2009).
(43) Hättig, C.; Weigend, F. *J. Chem. Phys.* **2000**, *113*, 5154−5161.
(44) Eichkorn, K.; Treutler, O.; Öhm, H.; Häser, M.; Ahlrichs, R. *Chem. Phys. Lett.* **1995**, *242*, 652−660.
(45) Eichkorn, K.; Weigend, F.; Treutler, O.; Ahlrichs, R. *Theor. Chem. Acc.* **1997**, *97*, 119−124.

(46) Semrouni, D.; Clavaguéra, C.; Dognon, J.-P.; Ohanessian, G. *Int. J. Mass Spectrom.* **2010**, *297*, 152−161.
(47) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *J. Chem. Phys.* **2010**, *132*, 154104.
(48) McQuarrie, D. A. *Statistical Mechanics*; Harper Collins Publishers: New York, 1976.
(49) Schmitz, M.; Tavan, P. *J. Chem. Phys.* **2004**, *121*, 12247−12258.
(50) Ponder, J. W. TINKER - Software Tools for Molecular Design, 2010. http://dasher.wustl.edu/tinker (accessed April 2011).
(51) Forbert, H.; Kleinschmidt, V. *Fourier*; Ruhr-Universität Bochum: 2002−2005 modified by A. Kohlmeyer, Version 21.02.2006.
(52) Borysow, J.; Moraldi, M.; Frommhold, L. *Mol. Phys.* **1985**, *56*, 913−922.
(53) Ramírez, R.; López-Ciudad, T.; P, P. K.; Marx, D. *J. Chem. Phys.* **2004**, *121*, 3973−3983.
(54) Marx, D.; Hutter, J. *Ab Initio Molecular Dynamics: Basic Theory and Advanced Methods*; Cambridge University Press: Cambridge, 2009.
(55) Marjolin, A.; Gourlaouen, C.; Clavaguéra, C.; Ren, P.; Wu, J.; Gresh, N.; Dognon, J.-P.; Piquemal, J.-P. *Theor. Chem. Acc.* **2012**, *131*, 1−14.
(56) Zhang, J.; Yang, W.; Piquemal, J.-P.; Ren, P. *J. Chem. Theory Comput.* **2012**, *8*, 1314−1324.
(57) Semrouni, D.; Isley, W. C. I.; Clavaguéra, C.; Dognon, J.-P.; Cramer, C. J.; Gagliardi, L. *J. Chem. Theory Comput.* **2013**, *9*, 3062−3071.
(58) Schnieders, M. J.; Fenn, T. D.; Pande, V. S. *J. Chem. Theory Comput.* **2011**, *7*, 1141−1156.
(59) Rasmussen, T. D.; Ren, P.; Ponder, J. W.; Jensen, F. *Int. J. Quantum Chem.* **2007**, *107*, 1390−1395.
(60) Jiao, D.; Zhang, J.; Duke, R. E.; Li, G.; Schnieders, M. J.; Ren, P. *J. Comput. Chem.* **2009**, *30*, 1701−1711.
(61) Schnieders, M. J.; Ponder, J. W. *J. Chem. Theory Comput.* **2007**, *3*, 2083−2097.
(62) Ren, P. Y.; Ponder, J. W. *J. Comput. Chem.* **2002**, *23*, 1497−1506.
(63) Zheng, X.; Wu, C.; Ponder, J. W.; Marshall, G. R. *J. Am. Chem. Soc.* **2012**, *134*, 15970−15978.
(64) Norton, J. E.; Brédas, J.-L. *J. Am. Chem. Soc.* **2008**, *130*, 12377−12384.
(65) Rasmussen, T. D.; Ren, P.; Ponder, J. W.; Jensen, F. *Int. J. Quantum Chem.* **2007**, *107*, 1390−1395.
(66) Cisneros, G. A.; Karttunen, M.; Ren, P.; Sagui, C. *Chem. Rev.* **2013**, *114*, 779−814.
(67) Tholé, B. *Chem. Phys.* **1981**, *59*, 341−350.
(68) Halgren, T. A. *J. Am. Chem. Soc.* **1992**, *114*, 7827−7843.
(69) Allinger, N. L.; Yuh, Y. H.; Lii, J. H. *J. Am. Chem. Soc.* **2002**, *111*, 8551−8566.
(70) Nevins, N.; Allinger, N. L. *J. Comput. Chem.* **1996**, *17*, 730−746.
(71) Halgren, T. A. *J. Comput. Chem.* **1996**, *17*, 553−586.
(72) Derreumaux, P.; Vergoten, G. *J. Chem. Phys.* **1995**, *102*, 8586−8605.
(73) Stone, A. J. *Chem. Phys. Lett.* **1981**, *83*, 233−239.
(74) Stone, A. J.; Alderton, M. *Mol. Phys.* **1985**, *56*, 1047−1064.
(75) Ren, P.; Ponder, J. W. *J. Comput. Chem.* **2002**, *23*, 1497−1506.
(76) Stone, A. J. *Distributed Multipole Analysis for Gaussian Wavefunctions*, version 2.2.03; 2005−2007.
(77) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09* Revision A.1; Gaussian Inc.: Wallingford, CT, 2009.
(78) Nosé, S. *J. Chem. Phys.* **1984**, *81*, 511−519.
(79) Hoover, W. G. *Phys. Rev. A* **1985**, *31*, 1695−1697.
(80) Oomens, J.; Sartakov, B. G.; Meijer, G.; von Helden, G. *Int. J. Mass Spectrom.* **2006**, *254*, 1−19.
(81) MacAleese, L.; Maître, P. *Mass Spectrom. Rev.* **2007**, *26*, 583−605.
(82) Parneix, P.; Basire, M.; Calvo, F. *J. Phys. Chem. A* **2013**, *117*, 3954−3959.
(83) Ataka, S.; Takeuchi, H.; Tasumi, M. *J. Mol. Struct.* **1984**, *113*, 147.
(84) Albrecht, M.; Rice, C. A.; Suhm, M. A. *J. Phys. Chem. A* **2008**, *112*, 7530−7542.
(85) Kubelka, J.; Keiderling, T. A. *J. Phys. Chem. A* **2001**, *105*, 10922−10928.
(86) Torii, H.; Tatsumi, T.; Kanazawa, T.; Tasumi, M. *J. Phys. Chem. B* **1998**, *102*, 309−314.
(87) The IRMPD spectrum of $(NMA)_2$-Na^+ was obtained at the Centre Laser Infrarouge dOrsay (CLIO) using a free electron laser (FEL) coupled to a 7T FT-ICR mass spectrometer. The extent of fragmentation is expressed as the fragmentation efficiency, F_{eff}, defined as $F_{eff} = -log[I_p/(I_p + \sum I_{frag})]$, in which I_p and I_{frag} are the parent and fragment ion intensities. Experimental parameters were very similar to those described in: Balaj, O. P.; Kapota, C.; Lemaire, J.; Ohanessian, G. *Int. J. Mass Spectrom.* **2008**, *269*, 196−209.
(88) Miller, D. J.; Lisy, J. M. *J. Phys. Chem. A* **2007**, *111*, 12409−12416.
(89) Jarrold, M. F. *Phys. Chem. Chem. Phys.* **2007**, *9*, 1659−1671.
(90) Stearns, J. A.; Boyarkin, O. V.; Rizzo, T. R. *J. Am. Chem. Soc.* **2007**, *129*, 13820−13821.
(91) Stearns, J. A.; Seaiby, C.; Boyarkin, O. V.; Rizzo, T. R. *Phys. Chem. Chem. Phys.* **2009**, *11*, 125−132.
(92) Albrieux, F.; Calvo, F.; Chirot, F.; Vorobyev, A.; Tsybin, Y. O.; Lepère, V.; Antoine, R.; Lemoine, J.; Dugourd, P. *J. Phys. Chem. A* **2010**, *114*, 6888−6896.
(93) Martens, J. K.; Compagnon, I.; Nicol, E.; McMahon, T. B.; Clavaguéra, C.; Ohanessian, G. *J. Phys. Chem. Lett.* **2012**, *3*, 3320−3324.
(94) Barlow, D. J.; Thornton, J. M. *J. Mol. Biol.* **1988**, *201*, 601−619.
(95) Sediki, A.; Snoek, L. C.; Gaigeot, M.-P. *Int. J. Mass Spectrom.* **2011**, *308*, 281−288.
(96) Rossi, M.; Scheffler, M.; Blum, V. *J. Phys. Chem. B* **2013**, *117*, 5574−5584.