Vibrational Spectra of Anhydrous and Monohydrated Caffeine and Theophylline Molecules and Crystals

Perla B. Balbuena,* Wendy Blocker, Rachel M. Dudek, Fredy A. Cabrales-Navarro, and Pussana Hirunsit

Department of Chemical Engineering, Texas A&M University, College Station, Texas 77843

Received: June 22, 2008; Revised Manuscript Received: August 29, 2008

Density functional theory and classical molecular dynamics simulations are used to investigate the vibrational spectra of caffeine and theophylline anhydrous and monohydrate molecules and those of their crystalline anhydrous and monohydrated states, with emphasis in the terahertz region of the spectra. To better understand the influence of water in the monohydrate crystal spectra, we analyze the vibrational spectra of water monomer, dimer, tetramer, and pentamer, and also those of liquid water at two different temperatures. In small water clusters, we observe the progressive addition of translational and librational modes to the terahertz region of the spectra. The water spectra predicted by rigid and flexible water models is examined with classical molecular dynamics, and the respective peaks, especially in the terahertz region, are compared with those found in the small clusters. Similar analysis done for caffeine and theophylline monohydrate molecules using density functional theory clearly shows the presence of water modes in the librational states and in the water stretching region. Molecular dynamics of caffeine and theophylline anhydrous and monohydrate crystals reveal the influence of vibrations from the molecule−molecule (caffeine or theophylline) crystal stacks and those from the water−molecule interactions found in the monohydrate molecules and new modes from molecule−molecule, water−molecule, and water−water hydrogen bonding interactions arising from collective effects in the crystal structure. Findings illustrate challenges of terahertz technology for the detection of specific substances in condensed phases.

1. Introduction

The technology of terahertz pulsed spectroscopy (TPS) is currently being developed to image tissues for the purpose of detecting skin cancer[1] and is expected to be extensively useful in sensors and biomedical applications where detection of individual chemical compounds or intermolecular interactions is crucial. This technology is good for imaging because it is nonionizing and sample nondestructive and by its nature excites vibrational frequencies arising due to van der Waals interactions and hydrogen bonds, making it an excellent technology for detecting water.[1] Terahertz radiation's ability to excite intermolecular modes in water also makes it a good candidate to monitor pharmaceutical purity. It also has recently been demonstrated to be a useful technique for the investigation of solid state properties in pharmaceutical materials,[2−5] where establishing the structure of a compound, particularly in its solid state, is an important factor to understand related physicochemical properties. Commonly these compounds can exist in different polymorph and hydrated forms.[6,7] The existence of these various conformers or pseudopolymorphs depends on a multitude of factors, from temperature and humidity to the process by which the compound is synthesized.[6,8] The presence of water in a molecular drug sample (often due to water being the solvent of crystallization) can cause it to adopt a completely different crystal structure (known as a pseudopolymorph) than that of the desired anhydrous structure.[7] As a consequence, physicochemical properties such as density, solubility, melting point, thermal conductivity, chemical and physical stability, and bioactivity may differ from the anhydrates.[1,7] These properties determine the effectiveness of a drug and its performance and could have impact at therapeutic, manufacturing, commercial, and legal levels.[7,9]

At present, there are thought to be three distinct classes of hydrate crystal structures, according to the distribution of water within the crystal lattice, channel, isolated-site, and ion-bound hydrates, each exhibiting different solid-state phase-transformation mechanisms.[3,10] These mechanisms and consequences thereof are currently not well understood,[3] and molecular modeling techniques can help elucidating the various structures and the phase-transformation mechanisms. Advantages of molecular modeling techniques as characterization tools to complement and guide experiments include a greater flexibility to investigate complicated geometries such as those present in biological molecules for comparison with experimental data and the possibility of obtaining theoretical predictions to interpret and describe spectroscopic data. Density functional theory (DFT) is commonly used to predict gas-phase spectra of isolated molecules, dimers, and clusters. In condensed phases, due to the complexity involved in periodic DFT modeling in large systems involving macromolecules, this remains a difficult problem.[3] Classical molecular dynamics (MD) simulation[11] is a complementary powerful tool where information about the vibrational spectra including intermolecular modes can be obtained by analysis of the Fourier transform of the velocity autocorrelation function. Although limited by the approximations of the effective force fields utilized to describe the intermolecular interactions, this technique is particularly useful for analysis of the terahertz domain, which is important for detection of specific molecules present in mixtures in condensed phases.

* To whom correspondence should be addressed. E-mail: balbuena@tamu.edu.

10.1021/jp805499m CCC: $40.75
Published on Web 09/25/2008

In this work, DFT and classical MD methods have been applied to generate vibrational spectra of water, caffeine ($C_8H_{10}N_4O_2$), and theophylline ($C_7H_8N_4O_2$). We studied anhydrous β-caffeine, caffeine monohydrate, anhydrous theophylline, and monohydrated theophylline. We have selected caffeine and theophylline, which are commonly used as medicines, as model molecules to investigate the effect of water in their vibrational spectra, both as isolated molecules and in the crystallized state. Calculated spectra are subsequently analyzed to identify vibrational modes and to investigate structural differences and intra- and intermolecular interactions that are responsible for variations in the positions of the calculated modes. The two different theoretical approaches utilized to determine the structures and vibrational spectra are complementary, and the ab initio analysis in small clusters greatly helps to elucidate the contributions to observed vibrational modes in condensed phases.

2. Methodology

Structures and molecular spectra of water monomer, dimer, tetramer, and pentamer structures were investigated using density functional theory (DFT) with Becke-3 (B3) and Perdew−Wang 91 (PW91) functionals[12,13] and the 6-311G, 6-311G(d,p), and 6-311G(d,p)++ basis sets. For caffeine and theophylline, the B3PW91/6-311G(d,p) method and basis set were used. Coordinates for the hydrated forms of caffeine were imported from the Cambridge Crystallographic database, reference code CAFINE01. The β-caffeine unit cell was only firmly established as of 2007, by Lehmann and Stowasser.[14] The structure was constructed and optimized in Cerius[2,15] based on representations of the cell given in ref 14. The structures for both anhydrous theophylline and the monohydrate were initially from the Cambridge Structural Database, BAPLOT01 and THEOPH01, respectively. These structures were optimized using DFT. All the ab initio calculations were done using the Gaussian 03[16] program.

MD simulations, which can describe larger systems incorporating collective as well as pressure and temperature effects were performed using DL_POLY 2.18.[17] The simulated systems include (1) water at 298.15 K at a density of 0.995 g/cm^3 and 374 K at a density of 0.95 g/cm^3 (both densities were obtained from coexistence data presented by Guissani and Guillot),[18] (2) a β-phase anhydrous caffeine and a caffeine monohydrate, both of approximate density of 1.43 g/cm^3, and (3) anhydrous and monohydrate theophylline. The *NVT* ensemble with Evans thermostat[19] was used for water, and the *NVT* ensemble with Hoover−Nose thermostat[20] was used for caffeine and theophylline. The Dreiding force field[21] was applied to represent the intra- and intermolecular interactions of caffeine and theophylline. Water was represented by the simple point charge extended (SPC/E) model,[22] where the bonds and bending angle are rigid and intermolecular forces are described by Lennard-Jones (LJ) and Coulombic terms, and also by flexible models. We follow the flexible water model introduced by Ferguson,[23] which contains bond stretching and angle bending terms and LJ and Coulombic terms for nonbonded interactions. We have tested two forms of the flexible model: one containing harmonic terms for both bonds and angles and another that includes a quadratic and a cubic term for bond stretching and a quadratic term for angle bending. Periodic boundary conditions in the three spatial directions were employed in all cases. The total simulation time was 300 ps, and the equilibration time was set at 200 ps. Longer simulations (1 ns) for the two forms of the caffeine crystal yielded negligible variations in the calculated spectra with respect to the shorter simulations. The time step was 0.001 ps. Fourier transforms of the velocity autocorrelation functions representing the vibrational spectra were computed and analyzed after the simulations were completed.

3. Results and Discussion

3.1. Water. ***3.1.1. DFT Results.*** Figure 1 shows the results of two basis sets, one including polarization and the other polarized and diffuse functions, on the IR water spectra, which can be divided into translations, librations, intramolecular bending, and symmetric and asymmetric stretching modes, approximately assigned to the 0−200, 200−1000, 1600−1800, 2900−3500, and 3500−3900 cm^{-1} regions, respectively. The three peaks on the calculated vibrational spectrum of a single water molecule using B3PW91 and basis sets 6-311G, 6-311G(d,p), and 6-311G(d,p)++ correspond to a bending mode at 1613/1636/1605 cm^{-1}, symmetric stretching at 3679/3846/3850 cm^{-1}, and asymmetric stretching at 3841/3947/3956 cm^{-1}. The three values listed for each mode correspond to each of the basis sets, respectively; in all cases, the calculated bending mode has the highest intensity peak. These can be compared with the experimental values of 1594.75, 3657.05, and 3755.93 cm^{-1} respectively.[24,25] However, we have not used scaling factors that are usually incorporated to correct the calculated vibrational frequencies.[26]

As we add a second water molecule to form a dimer, peaks begin to appear in the translations and librations region of the spectrum[27] as shown in Figure 1. In water clusters, the number of bending modes always equal the number of molecules in the system, and in general the total number of modes in each region increases, some at higher and some at lower frequencies than seen in the monomer. This is due to the particular symmetry of the water clusters and is especially notorious in the intermolecular region of the spectra.

Figure 2 shows water spectra in the terahertz region, a translations and librations region that includes rocking and rotation of the molecules and intermolecular interactions, specifically hydrogen bonding.[27] The spectra of water clusters of increasing number of molecules depend strongly on the specific geometry. There are many ways in which even a small number of water molecules can combine into a larger system. However, water−water interactions are dominated by hydrogen bond interactions, which limit the possible combinations to those giving minimum energy configurations.

Figure 3 shows the calculated dimer, tetramer, and pentamer in the conformations corresponding to their respective minimum energy. The three basis sets gave similar structures, which are provided as Supporting Information. The dimer forms one H-bond structure with one molecule acting as a donor and the other as an acceptor. Keutsch and Saykally described three main transitions in the dimer structure: the first with a barrier of ∼157 cm^{-1} corresponding to the exchange of the protons in the acceptor monomer and rotation of the donor monomer, a second one with a barrier of ∼207 cm^{-1} that allows an exchange of the roles of the hydrogen bond donating and accepting monomers, and a third barrier of ∼394 cm^{-1} that combines librational motion of the donor with flipping motion of the acceptor. These three translational/rotational motions could correspond to the three first peaks in Figure 2. In addition, there is a peak at 700 cm^{-1}, which corresponds to librations.

A tetramer with the oxygen atoms forming the four corners of a highly symmetric cyclic structure is the result of the structure tending to maximize the number of hydrogen bonds. There are two hydrogen bonds per molecule, and the structure could be defined as up−down−up−down, in relation to the

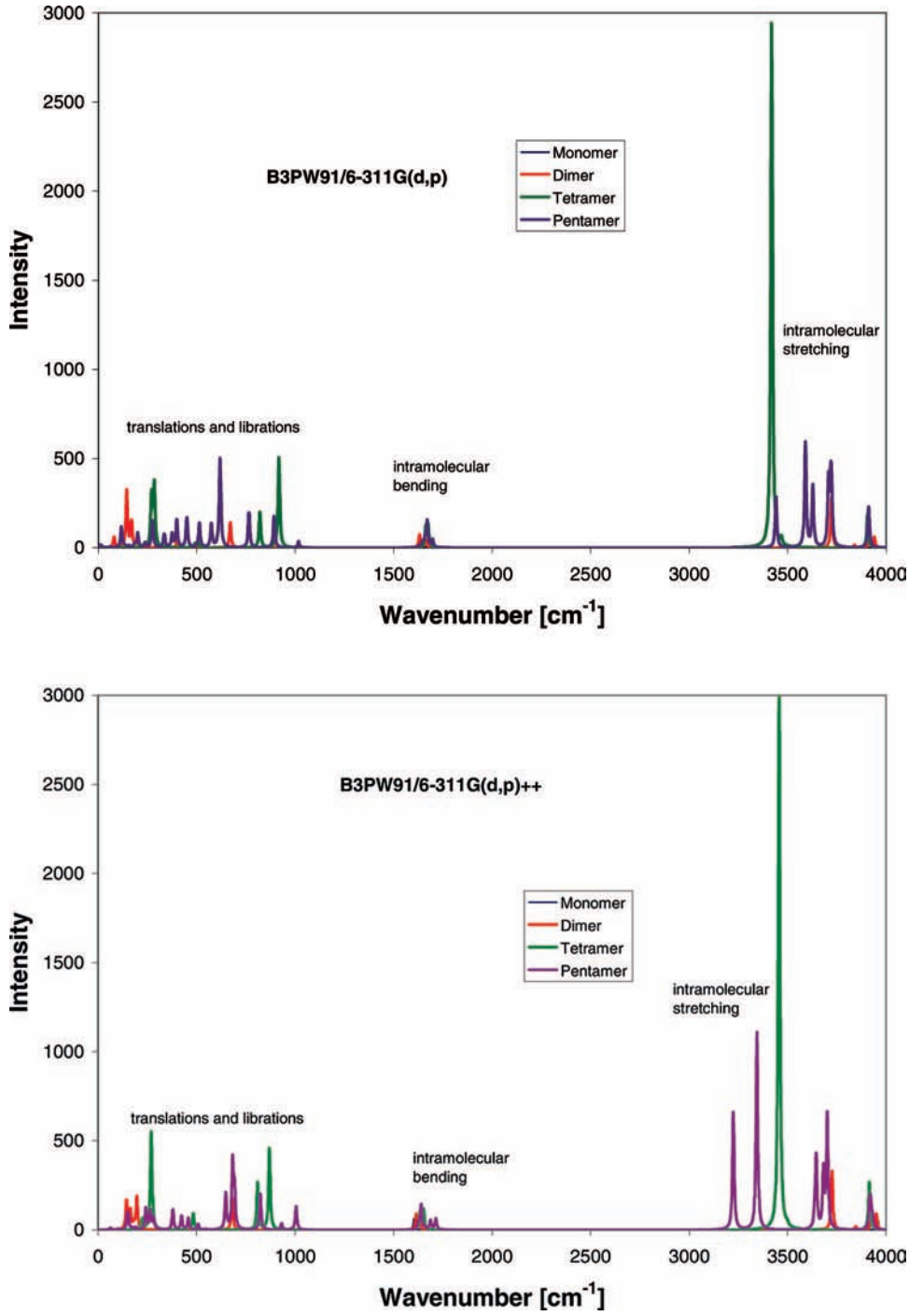

Figure 1. B3PW91/6-311G(d,p) (top) and B3PW91/6-311G(d,p)++ (bottom) calculated water spectra for $(H_2O)_n$, with $n = 1$, 2, 4, and 5.

position of the free hydrogens in each monomer.[27] This cyclic structure creates a unique spectrum where the highest intensity peak occurs at 3419/3456 cm^{-1} (Figure 1) for 6-311G(d,p) and 6-311G(d,p)++, respectively, involving the symmetric stretching of the hydrogen bonds of two opposite oxygen atoms, sharing the hydrogen atoms back and forth. There is a small intensity peak at a frequency below 100 cm^{-1} that corresponds to translational in-plane ring deformation, and the well-defined peaks at ~300 and 800−900 cm^{-1} (Figure 2) reflect coordinated rotational−vibrational motions.

Finally, a last water molecule is added to form a pentamer where the modes are a combination of those found in the dimer and in the tretramer. In the optimized pentamer (Figure 3), there are two molecules participating in three H-bonds each, and the other three participate in two. Table 1 shows how the addition of water molecules to the structure successively increases the average length of the covalent bonds between the oxygen and hydrogen atoms, except in the pentamer because of the larger number of H-bonds. The hydrogen bond length is optimized in the tetramer, whereas it is elongated in the pentamer and dimer, and the HOH angle tends to increase with increasing number of water molecules in the system. Binding energies per water molecule are calculated using the difference of energies between the energy of the complex and n times the energy of a water molecule. For example, $BE = [E(H_2O)_n - nE(H_2O)]/n$, where BE is the binding energy, $E(H_2O)_n$ is the energy of the water cluster, $E(H_2O)$ is the energy of the water monomer, and n is the number of water molecules in the cluster; all the energies

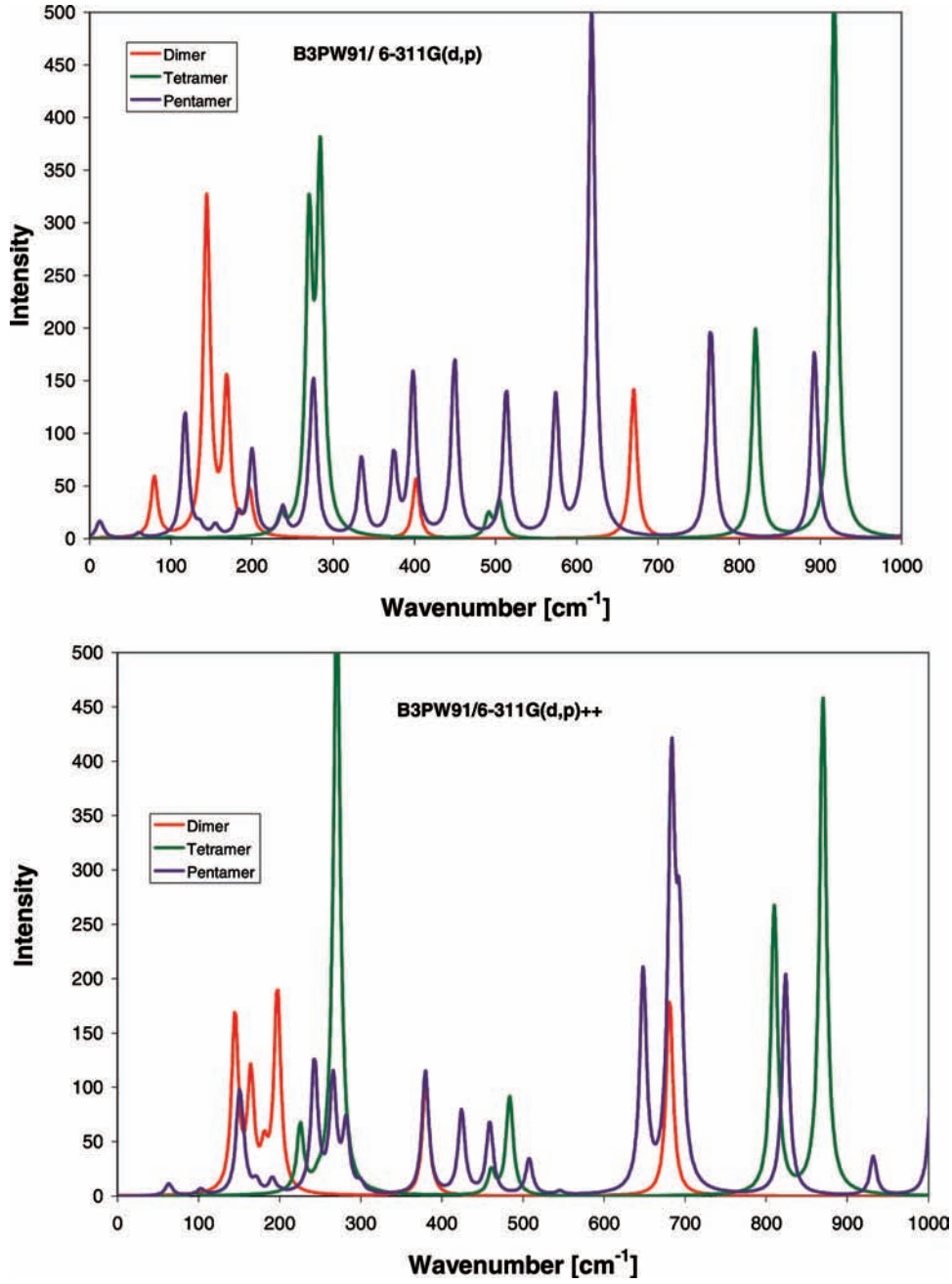

Figure 2. B3PW91 calculated water spectra (top, 6-311G(d,p); bottom, 6-311G(d,p)++) in the terahertz region showing intermolecular modes. Largest discrepancies between the results of the two basis sets in the librations region of the pentamer are due to differences in the optimized geometries.

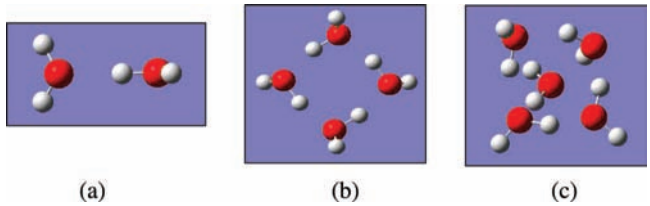

Figure 3. Optimized structures (B3PW91/6-311G(d,p)) of (a) dimer, (b) tetramer, and (c) pentamer. Results from both basis sets with spectra reported in Figures 1 and 2 yield similar structures for dimer and tetramer; the pentamer structure differs between the results of the two basis sets. Strong hydrogen bond networks are observed in all cases.

are corrected by adding the zero point energy. They are summarized in Table 1. Both calculations reveal the optimum configuration of the tetramer maximizing the binding energy per molecule. The binding energy results of B3PW91/6-311G(d,p) for the dimer (−4.61 kcal/mol) are in good agreement with that reported (−4.85 kcal/mol) by Keutsch and Saykally.[27] We have also calculated the binding energy given by the functional M05-2X,[28] which is expected to describe more adequately weak dispersion interactions including hydrogen bonding. The results are −5.15 kcal/mol using 6-311G(d,p) and −3.90 kcal/mol adding diffuse functions, 6-311G(d,p)++, which together with the binding energies for the dimer shown in Table 1 suggest that the addition of diffuse functions tends to underestimate the water binding energy.

3.1.2. MD Results. A water system density of 0.995 g/cm^3 was run at room temperature, 300 K, and another at an elevated temperature of 374 K and at a density of 0.95 g/cm^3, using both the rigid SPC/E model and the flexible models described in the Methodology section. The two flexible models predict identical

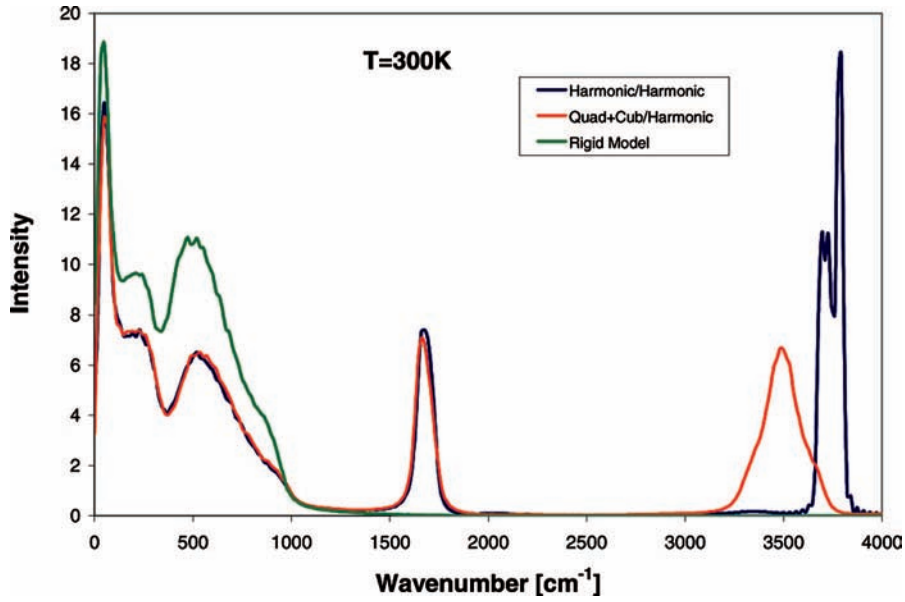

Figure 4. Vibrational spectra of liquid water at 300 K obtained by Fourier transform of the velocity autocorrelation functions from MD simulations using the SPC/E rigid model (only reveals intermolecular interactions) and two different flexible water models.

TABLE 1: Average Bond Lengths, Angles, And Binding Energies for Calculated Water Clusters (B3PW91 and the Basis Set Indicated)

	covalent OH bond (Å)		hydrogen bond (Å)		HOH angle (deg)		binding energy per water molecule (kcal/mol)	
system	6-311G (d,p)	6-311G (d,p)++	6-311G (d,p)	6-311G (d,p)++	6-311G (d,p)	6-311G (d,p)++	6-311G (d,p)	6-311G (d,p)++
monomer	0.960	0.960			103.8	104.9		
dimer	0.962	0.962	1.92	1.93	104.2	105.1	−4.61	−2.97
tetramer	0.973	0.972	1.73	1.75	105.0	106.0	−7.16	−5.05
pentamer	0.969	0.970	1.89	1.86	105.2	105.9	−5.58	−4.64

spectra in the terahertz region (Figure 4), but they show a substantial difference in the OH stretching intramolecular region. As remarked by Ferguson,[23] the pure harmonic potential (harmonic for bonds and angle) fails to yield a red shift in the stretching compared with the gas-phase spectrum, yielding peaks between 3500 and 4000 cm^{-1} very similar to those shown in Figure 1 for the water clusters. The other flexible model, which includes quadratic and cubic terms for the bonds and only a quadratic term for the angle indeed shows the red shift for the stretching modes. In the terahertz region, the rigid and flexible models agree qualitatively. There are three well-defined regions: at about 50 cm^{-1}, there is a sharp peak that we attribute to translational motion (OOO bending) experimentally detected at 65 cm^{-1};[29] at about 200 cm^{-1}, the calculated peak is another translational motion (OO stretching) experimentally found at 162 cm^{-1};[29] and the third region (600−1000 cm^{-1}) includes librational motions. The rigid model enhances all intermolecular modes in comparison to the flexible models, the molecular rigidity allows for higher intensities in the modes, which correspond to intermolecular OO stretching and OOO bending.

The MD calculated intensities for the vibrational peaks cannot be compared with those from DFT spectra due to the different nature of the calculations; however, peaks positions and relative intensities can be compared. Looking at the spectra of the individual water clusters, the pentamer's spectrum shows the largest similarities to the MD spectra. The dimer has too few intermolecular interactions and does not compare well. The tetramer is a special case because of the cyclic structure formed that amplified some of its peaks, making it hard to compare.

The structure of the pentamer, however, has all the molecules evenly spaced in a conformation that easily allows other water molecules to bind with it. However, we note that the bulk system may contain contributions from pair, three-body, and many-body interactions,[27] and therefore, an imaginary envelope involving all modes from the small clusters may be the best representation showing the tendency to converge toward the larger system. Such an imaginary envelope (in Figure 2) could be correlated to modes observed in the MD water vibrational spectra (Figure 4) as follows: a low intensity peak around 60−80 cm^{-1} and others in the region ∼160−200 cm^{-1} in the translational region, peaks from around 220−300 cm^{-1} in the H-bond rocking and librations region, and peaks within the 400−700 cm^{-1} range in the librations region (Figure 2), all similar to the MD spectra (Figure 4). There is a final group of peaks in the DFT spectrum (800−1000 cm^{-1}) that may have been converted in the shoulder seen in Figure 4 at the highest frequencies.

According to Figure 5, the higher temperature system tends to have broader peaks shifted to lower frequencies. An increase in temperature causes higher molecular diffusion and a decrease in the lifetimes[30] of hydrogen bonds; the broadened peaks in the elevated temperature spectrum illustrate this diffusion. Similar behavior was found for the other flexible model and for the rigid model (not shown). Carey and Korenowski[29] have reported the temperature dependence of the Raman modes in water as a function of temperature. They show that the translational and librational modes experience a clear red shift as temperature increases from 22 to 400 °C, in good qualitative agreement with the MD results in Figure 5. The radial distribution function (RDF) of the molecules was also graphed, not shown here. The RDF graphs show that the intermolecular bond distances do not change significantly with temperature.

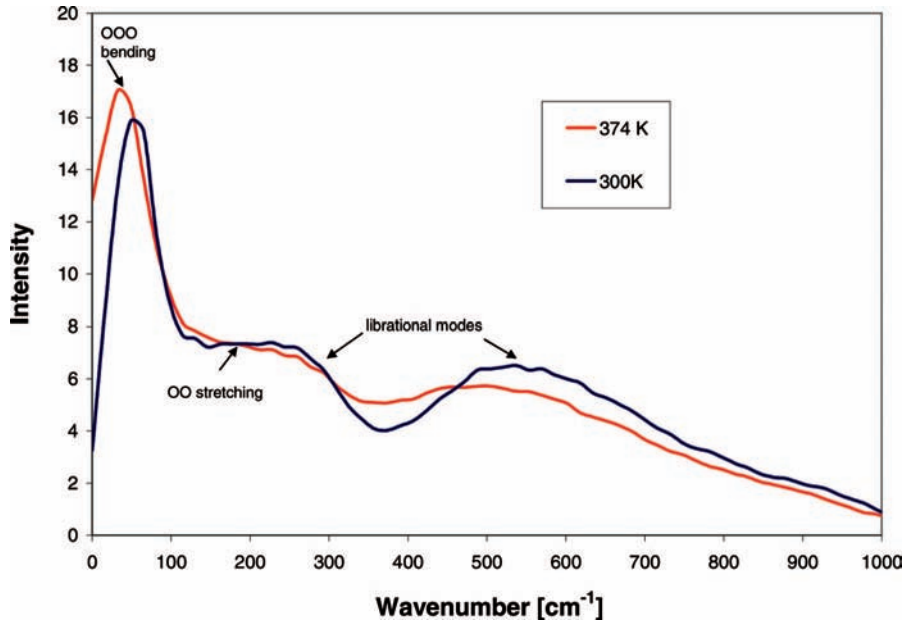

Figure 5. Vibrational spectra of liquid water obtained by Fourier transform of the velocity autocorrelation functions from MD simulations using a flexible model[23] including quadratic and cubic terms for the bonds and quadratic terms for the angles. The spectra show collective effects at two different temperatures. The decrease in the lifetimes of hydrogen bonds at the highest temperature broadens the peaks and shifts the maxima to lower frequencies.

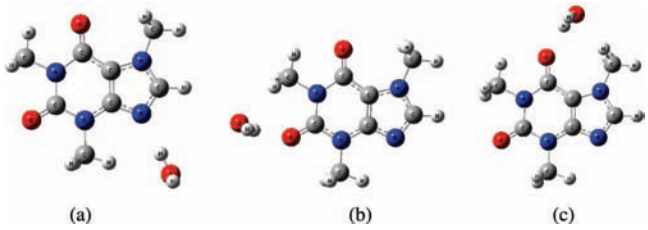

Figure 6. B3PW91/6-311G(d,p) optimized geometries of caffeine monohydrates: (a) monohydrate-1 where one of the H atoms of water interacts with a N atom (N···H distance 1.96 Å); (b) monohydrate-2 where one of the H atoms of water interacts with one of the caffeine O atoms (O···H distance 1.95 Å); (c) monohydrate-3 where one of the H atoms of water interacts with another of the caffeine O atoms (O···H distance 1.91 Å). O atoms are red, N atoms blue, C atoms gray, and H atoms white. Structures are provided as Supporting Information.

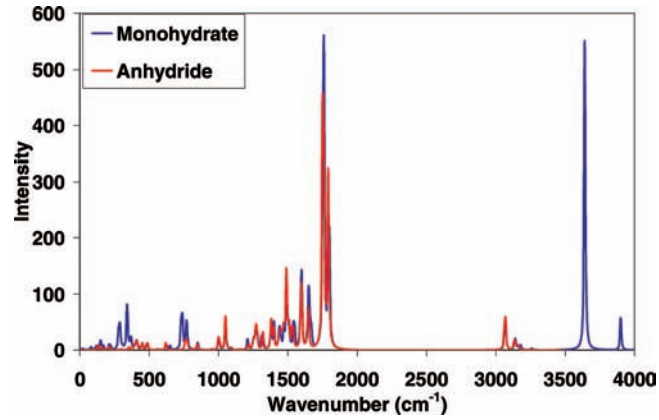

Figure 7. Overlay plot of caffeine β-anhydride and caffeine monohydrate IR spectra obtained from DFT calculations of a single caffeine molecule system (red) and a one-caffeine−one-water molecule (blue).

However, there are fewer of these interactions at the higher temperature. Particularly the number and lifetimes of hydrogen bonds, which are the main interactions that give liquid water its typical structure, decrease as the temperature increases due to the higher kinetic energy of the molecules.[31]

3.2. Caffeine. ***3.2.1. DFT Results.*** After optimization of a single caffeine molecule, initial configurations of caffeine monohydrates were constructed with the water molecule positioned to bind to three possible sites as shown in Figure 6. The calculated binding energies (B3PW91/6-311G(d,p)) for the formation of the monohydrates are −5.74, −5.06, and −6.21 kcal/mol for monohydrates 1, 2, and 3, respectively.

Figure 7 shows the spectra of caffeine and its monohydrate-1 (structure of Figure 6a). From 1000 to 3200 cm^{-1}, the caffeine monohydrate and anhydride IR spectra from DFT are nearly superimposable. Monohydrate spectra exhibited some offset in the intermolecular region as well as in the 3600−3700 cm^{-1} (water stretching) region. Specifically, the modes of the monohydrate peak at 341 cm^{-1} are due to librations in the water molecule coupled to bending of the caffeine molecular plane, whereas the modes peaking at 735 cm^{-1} are due to coupled librations of water and caffeine.

This causes the spectrum in water stretching mode of monohydrate 1 to shift toward lower frequency with higher intensity than monohydrates 2 and 3 (Figure 8). Also, the librational mode described for monohydrate-1 at 735 cm^{-1} becomes shifted to lower frequencies (∼612 cm^{-1}) for monohydrates 2 and 3. The IR spectra of monohydrate 2 and 3 (both having intermolecular H_w-O_c interactions) do not show any significant differences.

3.2.2. MD Results. MD vibrational spectra for β-caffeine and monohydrate are shown in Figure 9 for the terahertz region only (0−1000 cm^{-1}) to highlight differences in vibrational frequencies and intensities between the two crystal structures. Above ∼650 cm^{-1}, with few exceptions, the peaks appear at the same frequencies and differ only in intensity. In the range between 57−250 cm^{-1}, we observe the most significant differences between vibrational spectra of β-caffeine and monohydrate. β-caffeine shows three peaks at 65.10, 113.93, and 301.11 cm^{-1} and two shoulders at 154.62 and 268.55 cm^{-1}, whereas caffeine monohydrate shows peaks in this range shifting to either higher or lower frequencies. Furthermore, they do not show peaks at

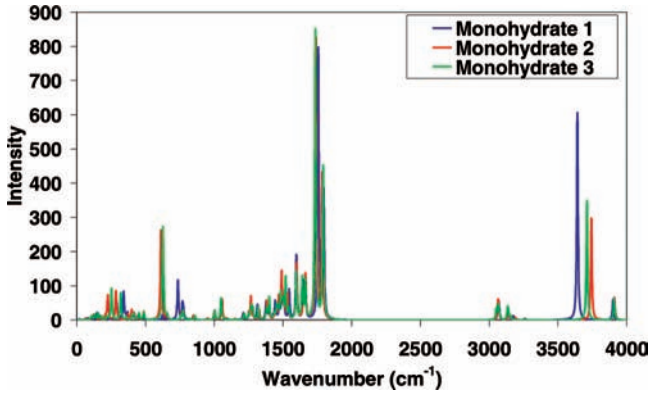

Figure 8. Overlay plot of IR spectra corresponding to caffeine monohydrates 1, 2, and 3 (geometries refer to Figure 6). Librational modes at ~700 cm^{-1} for monohydrate-1 where caffeine and water interact via H_wN_c bonds differ from the other two structures having H_wO_c intermolecular interactions. The other difference is in the water intramolecular OH stretching.

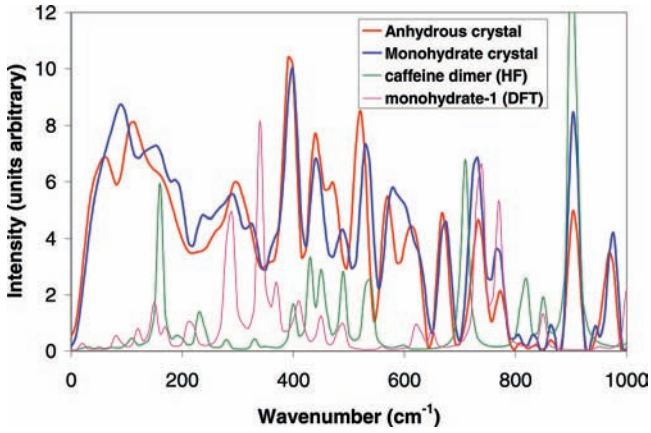

Figure 9. Overlay plot of caffeine monohydrate and anhydrous β-caffeine crystal vibrational spectra from MD. The red curve corresponds to the IR spectrum of a stacked caffeine dimer calculated with HF/6-31G and the pink curve to that of the caffeine monohydrate-1 (Figure 6a) calculated with B3PW91/6-311G(d,p); in both cases the calculated intensities have been reduced by 10 to improve visualization.

the same frequencies in this frequency range. To elucidate the modes of anhydrous caffeine crystal, we also calculated a stacked caffeine dimer using Hartree−Fock/6-31G (structure provided as Supporting Information) and determined its IR spectrum. The results (Figure 9) indicate the presence of weak intermolecular bending modes below 100 cm^{-1}, translational modes at ~160 cm^{-1} that resulted from a concerted motion of the methyl groups up and down the molecular planes. These motions could correspond to the first two peaks in the anhydrous caffeine crystal. Between 200 and 300 cm^{-1}, the stacked dimer shows in-plane bending motion of one of the molecules, and the group of peaks between 400 and 600 cm^{-1} are the result of in-plane bending and ring stretching, including a combined rocking motion of both molecules, which correspond to the ones observed in the crystal spectra. Other cooperative motions are the peaks at 710 and 903 cm^{-1} (the highest intensity peak of the caffeine dimer in the terahertz region), which correspond to concerted rocking−breathing motions of both molecules in the dimer. The matching of these dimer peaks with the ones of the anhydrous and hydrated crystal structures allow us to infer that such synchronized motions will exist also in the condensed phase and their intensity is enhanced by the presence of the water molecules in

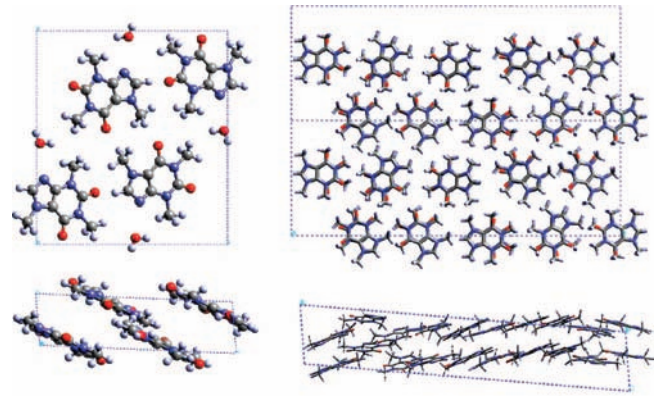

Figure 10. Caffeine monohydrate crystal cell (top left) and side view (bottom left) and anhydrous β-caffeine crystal cell (top right) and side view (bottom right).

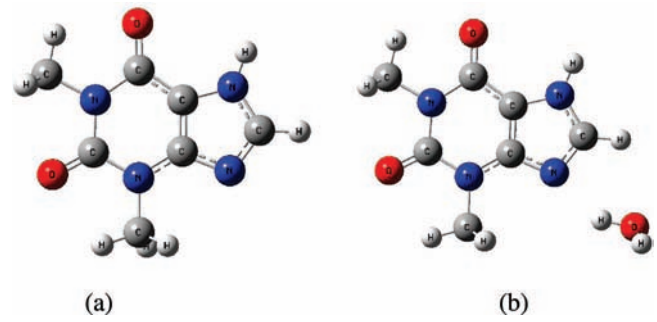

Figure 11. B3PW91/6-311G (d,p) optimized geometries of (a) theophylline and (b) theophylline monohydrate. The distance N···H in the monohydrate is 1.68 Å, and the NHO angle is 148.7°.

the monohydrate crystal. The IR spectrum of the caffeine monohydrate-1 molecule is also shown in Figure 9.

It is interesting to see that except for the first two peaks at frequencies <100 cm^{-1}, the other spectral features may be associated with either vibrational motions of the monohydrate molecule or those of the stacked caffeine dimer, and precisely those regions where we do not see any correspondence are the ones where the monohydrate and the hydrated crystal spectra differ the most. Although our current results cannot provide detailed information about the nature of these peaks, it is expected that caffeine−caffeine, caffeine−water, and water−water H-bonding interactions are involved. Visualizations of the caffeine monohydrate and β-phase anhydrous caffeine unit cells are shown in Figure 10. The monohydrate crystal structure corresponds to the case of monohydrate-1, where water forms a H-bond with the N atom of the imidazole group (Figure 6a), such interaction becomes weakened at 298 K, with an average distance N···H of 2.6 Å, compared with 1.96 Å calculated in the gas phase. However, new H-bonding interactions result from the collective ensemble.

3.3. Theophylline. ***3.3.1. DFT results.*** The theophylline molecule was optimized first followed by optimization of a number of different monohydrate configurations. Theophylline and one of the theophylline monohydrate optimized geometries are shown in Figure 11. Details are provided as Supporting Information.

The theophylline tetramer was also optimized, but due to the larger number of molecules in the system, a simpler level of theory was required. The method used was AM1, a semiempirical calculation.[32] The monomer and tetramer spectra (both from AM1) shown in Figure 12 reveal a great deal of similarity between them; the peaks are basically at the same positions,

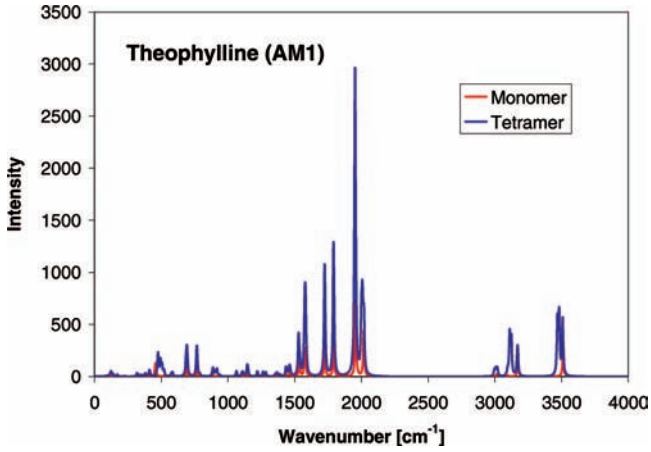

Figure 12. Calculated (AM1) spectra of theophylline monomer and tetramer.

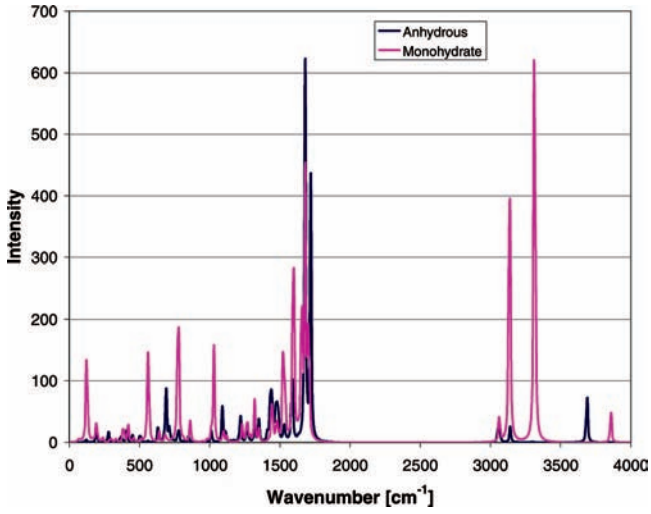

Figure 13. B3PW91/6-311G calculated spectra of theophylline monomer and monohydrate molecular structures shown in Figure 11.

but the intensities are enhanced on the tetramer. Additional intermolecular modes are shown in the terahertz region for the tetramer.

The spectrum of the theophylline monohydrate molecule (Figure 13) has some similarities to that of the anhydrate with variations due to interactions with the water molecules in the terahertz region and to the water intramolecular modes. Specifically, the monomer has weak in-plane bending modes at 200 cm^{-1} and ring stretching modes at 400 cm^{-1} and librational modes at 630 cm^{-1}, especially from the imidazole group, which is less restrained than that in caffeine due to the absence of one methyl group. At 1000−1500 cm^{-1}, we observe in-plane stretching and librational motion with increasing bending of the H atoms in the methyl groups; in the range 1600−1700 cm^{-1}, there is an increase of vibrations from C−C stretching, with C atoms moving in and out of the molecular plane, especially in the six-member ring, and at 3690 cm^{-1}, the NH stretching mode is detected.

On the other hand, the monohydrate molecule exhibits much higher motion in the terahertz region; at 123 cm^{-1} there is an in-plane bending coupled to water libration, at 153 cm^{-1} water rotation coupled to in-plane breathing motion, at 776−1000 cm^{-1} water libration involving N···H···O bonding and O···H···O interactions, and at 1600−1700 cm^{-1} in-plane stretching including C−C and C═O combined with similar

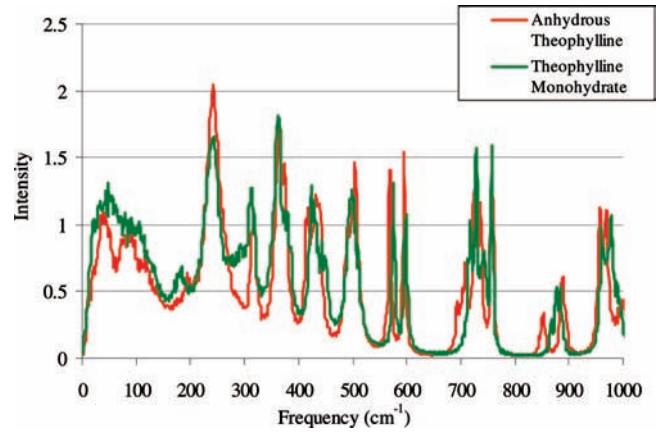

Figure 14. Molecular dynamics calculated spectra of anhydrous and monohydrate theophylline crystals in the terahertz region. Note that the most important differences come from the water−theophylline hydrogen bonding interactions in the low-frequency region of the spectrum.

H-bonding motions and water bending. At 3008 cm^{-1} we observe CH symmetric stretching, at 3313 cm^{-1} water OH symmetric stretching along with NH and CH stretchings, and at 3857 cm^{-1} water asymmetric stretching.

3.3.2. MD Results. Crystal structures of anhydrous and monohydrate theophylline were simulated at 298.15 K. As observed with caffeine, the largest differences in the spectra occur at frequencies below 1000 cm^{-1}. As shown by Figure 14, the spectrum of the monohydrate has higher intensities in the peak ranges 20−75, 160−190, 270−315, and 720−800 cm^{-1}. This increase in intensity is most likely due to the interactions of theophylline with the water molecules. In the simulation cell of the monohydrate representing the crystal structure (Figure 15), it is observed that the water molecules are actually close enough to interact with each other; comparisons between the spectrum of the water dimer and theophylline monohydrate crystal illustrate this occurrence.

These interactions contribute to the differences between the spectra. The water pentamer, as well as the water's MD spectra, has peaks in the 25−85 cm^{-1} range. This peak is also present in the DFT spectra for the monohydrate crystal but is absent in that of the anhydrate crystal and tetramer. Therefore, the presence of water in the monohydrate creates this difference in the spectra. The water−water interaction causes the peak at 160−190 cm^{-1} in the MD spectrum; this peak is also present in the water dimer spectra. The higher intensities at 270−315 cm^{-1}, can be compared with the peaks for the water tetramer and pentamer, again emphasizing the influence of water on the crystal structure of the hydrated theophylline.

The structural differences between the anhydrate and monohydrate molecules are significant and could also be responsible for differences in the spectra. Without water, theophylline has three hydrogen accepting sites but only two donating sites. The addition of the water molecule stabilizes this imbalance causing the monohydrate to be the thermodynamically favorable form. From the analysis done for the caffeine dimer, we infer that similar theophylline−theophylline interactions will dominate both terahertz spectra of the anhydrous and monohydrate forms. However, at low frequencies, we attribute the indicated differences to the water−water and water−theophylline interactions and those at ∼700 and ∼850 cm^{-1} to the water−theophylline modes as discussed for the theophylline monohydrate molecule. Visualizations of the unit cell for the anhydrate and monohydrate crystal as well as some geometric parameters obtained from the MD simulations are shown in Figure 15.

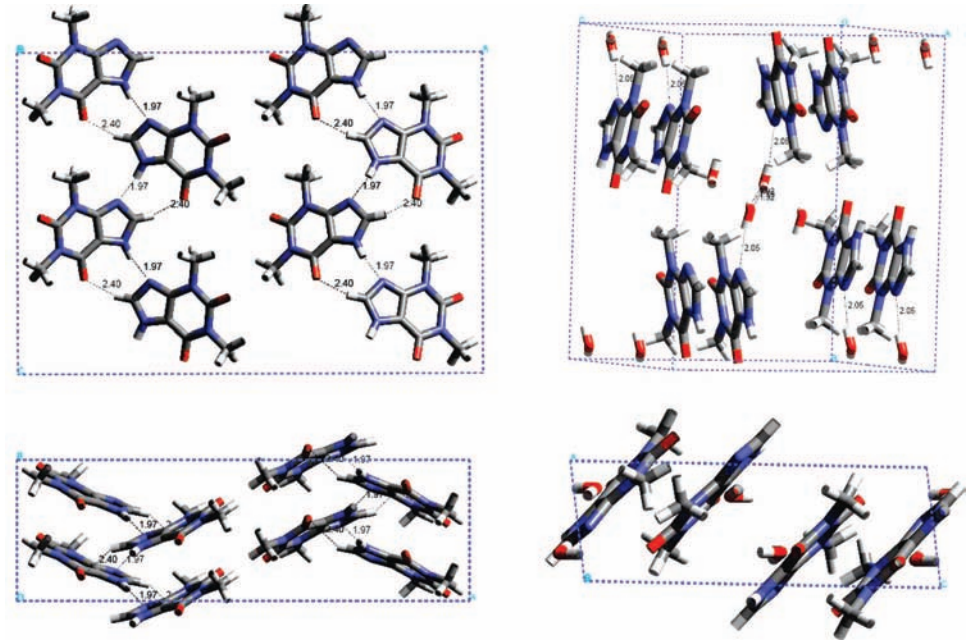

Figure 15. Snapshots of the unit cell for the anhydrate crystal (top left) and side view (bottom left) and monohydrate crystal (top right) and side view (bottom right). The main intermolecular distances are indicated (in Å). Note the significant elongation of the H-bonding N···H distances (~2.06 Å; these are instantaneous values) with respect to those in gas phase (Figure 11).

4. Conclusions

Combined quantum-mechanical and classical dynamics approaches illustrate the feasibility of analyses of spectroscopic data for the detection of certain species. It is clear that the terahertz region is dominated by modes arising from intermolecular interactions, although some contributions from intramolecular modes can be important especially with large molecules. Among those intermolecular interactions H-bonding and dispersion interactions are very commonly found in water and organic molecules, and disentangling those effects should be the first target for a terahertz-based spectroscopic analysis. Here we investigated two molecules, caffeine and theophylline, and we analyzed the single molecules, the monohydrates, and the condensed crystalline phases of the anhydrous and hydrated forms. In order to understand the condensed-phase spectra, particularly in the terahertz region, we found that it is useful to examine the structures and vibrational modes of the anhydrous molecular dimer and the monohydrate and the structure (H-bond formation) of the condensed phase. For example, the crystal spectra of anhydrous caffeine in the terahertz region shows clearly the contributions of intramolecular modes of caffeine as found in the single-molecule caffeine−caffeine interactions as detected in the dimer particularly contributing in the translational and librational region, whereas hydrated caffeine also has very similar vibrational modes but also exhibits water−caffeine modes as the ones observed in the monohydrate molecule. On the other hand, the regions of both spectra that cannot be associated with any of the individual species or small clusters must be due to H-bonding or dispersion interactions that arise because of the collective arrangement of molecules in the crystal. More work is needed to further elucidate these additional modes.

We remark that combining various levels of theory is inevitable at this stage of the computer software/hardware technology. The prediction of correct vibrational and rotational modes is tightly associated with obtaining the correct geometry. It has been demonstrated, for example, that Hartree−Fock methods predict good structures and also spectra in good agreement with the experiments; this agreement is as good as that obtained with DFT methods.[33] Semiempirical methods were used in this work to compare the monomer and tetramer structures of theophylline because of the high cost of computing the tetramer with higher level methods; the predicted structure of the semiempirical methods is less reliable; that is why we did not compare these results to those of DFT. Finally, we recognize that some modes, especially the intermolecular ones, are affected by temperature; this is one of the points that we show here, but given the good accuracy of the ab initio methods, they are certainly useful to assess whether a peak or a group of peaks may belong to a certain motion as can be inferred very well from molecules or small groups of molecules.

Acknowledgment. We gratefully acknowledge financial support (to W.B. and R.M.D.) from the National Science Foundation REU program, Grant Number 0552655, and from the US Army Research Office and the US Defense Threat Reduction Agency (DTRA). P.H. thanks the National Nanotechnology Center of Thailand for providing a scholarship.

Supporting Information Available: Structures of the water dimer, tetramer, and pentamer, calculated with B3PW91 and 6-311G, 6-311G(d,p), and 6-311G(d,p)++, of caffeine and caffeine monohydrates, of theophylline and theophylline monohydrate (B3PW91/6-311G(d,p)), of caffeine dimer (HF/6-31G), of theophylline, and of theophylline tetramer (AM1). This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Pickwell, E.; Wallace, V. P. Biomedical applications of terahertz technology. *J. Phys. D: Appl. Phys.* **2006**, R301–310.

(2) Zeitler, J. A.; Newnham, D. A.; Taday, P. F.; Threlfall, T. L.; Lancaster, R. W.; Berg, R. W.; Strachan, C. J.; Pepper, M.; Gordon, K. C.; Rades, T. Characterization of temperature-induced phase transitions in five polymorphic forms of sulfathiazole by terahertz pulsed spectroscopy and differential scanning calorimetry. *J. Pharm. Sci.* **2006**, *95*, 2486–2498.

(3) Zeitler, J. A.; Kogermann, K.; Rantanen, J.; Rades, T.; Taday, P. F.; Pepper, M.; Aaltonen, J.; Strachan, C. J. Drug hydrate systems and

dehydration processes studied by terahertz pulsed spectroscopy. *Int. J. Pharm.* **2007**, *334*, 78–84.

(4) Taday, P. F.; Bradley, I. V.; Arnone, D. D.; Pepper, M. Using terahertz pulse spectroscopy to study the crystalline structure of a drug: A case study of the polymorphs of ranitidine hydrochloride. *J. Pharm. Sci.* **2003**, *92*, 831–838.

(5) Strachan, C. J.; Rades, T.; Newnham, D. A.; Gordon, K. C.; Pepper, M.; Taday, P. F. Using terahertz pulsed spectroscopy to study crystallinity of pharmaceutical materials. *Chem. Phys. Lett.* **2004**, *390*, 20–24.

(6) Threlfall, T. L. Analysis of organic polymorphs—a review. *Analyst* **1995**, *120*, 2435–2460.

(7) Liu, H.-B.; Chen, Y.; Zhang, X. C. Characterization of anhydrous and hydrated pharmaceutical materials with THz time-domain spectroscopy. *J. Pharm. Sci.* **2007**, *96*, 927–934.

(8) Airaksinen, S.; Karjalainen, M.; Shevchenko, A.; Westermarck, S.; Leppänen, E.; Rantanen, J.; Yliruusi, J. Role of water in the physical stability of solid dosage formulations. *J. Pharm. Sci.* **2005**, *94*, 2147–2165.

(9) Byrn, S.; Pfeiffer, R.; Ganey, M.; Hoiberg, C.; Poochikian, G. Pharmaceutical Solids: A Strategic Approach to Regulatory Considerations. *Pharm. Res.* **1995**, *12*, 945–954.

(10) Morris, K. R. *Polymorphism in Pharmaceutical Solids*; Marcel Dekker Inc.: New York, 1999.

(11) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Oxford University Press: Oxford, U.K., 1990.

(12) Perdew, J. P. Unified theory of exchange and correlation beyond the local density approximation. In *Electronic Structure of Solids*; Ziesche, P., Eschrig, H., Eds; Akademie Verlag: Berlin, 1991.

(13) Perdew, J. P.; Wang, Y. Accurate and simple analytic representation of the electron-gas correlation energy. *Phys. Rev. B* **1992**, *45*, 13244–13249.

(14) Lehmann, Christian, W.; Stowasser, F. The crystal structure of anhydrous β-caffeine as determined from X-ray powder-diffraction data. *Chem.—Eur. J.* **2007**, *13*, 2908–2911.

(15) Molecular Simulations, I. *CERIUS 2*; San Diego, CA, 1997.

(16) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*, Revision C.02; Gaussian, Inc.: Wallingford, CT, 2004.

(17) Smith, W.; Forester, T. R. DL_POLY_2.0: A general-purpose parallel molecular dynamics simulation package. *J. Mol. Graphics* **1996**, *14*, 136–141.

(18) Guissani, Y.; Guillot, B. A computer-simulation study of the liquid-vapor coexistence curve of water. *J. Chem. Phys.* **1993**, *98*, 8221–8235.

(19) Evans, D. J.; Morriss, G. P. The isothermal/isobaric molecular dynamics ensemble. *Phys. Lett.* **1983**, *98A*, 433–436.

(20) Hoover, W. Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A* **1985**, *31*, 1695.

(21) Mayo, S. L.; Olafson, B. D.; Goddard, W. A. Dreiding: A generic force field for molecular simulations. *J. Phys. Chem.* **1990**, *94*, 8897–8909.

(22) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. The missing term in effective pair potentials. *J. Phys. Chem.* **1987**, *91*, 6269–6271.

(23) Ferguson, D. M. Parameterization and evaluation of a flexible model of water. *J. Comput. Chem.* **1995**, *16*, 501–511.

(24) Bernath, P. F. The spectroscopy of water vapour: Experiment, theory and applications. *Phys. Chem. Chem. Phys.* **2002**, *4*, 1501–1509.

(25) Lemus, R. Vibrational excitations in H_2O in the framework of a local model. *J. Mol. Spectrosc.* **2004**, *225*, 73–92.

(26) Scott, A. P.; Radom, L. Harmonic vibrational frequencies: An evaluation of Hartree−Fock, Moller−Plesset, quadratic configuration interaction, density functional theory, and semiempirical scale factors. *J. Phys. Chem.* **1996**, *100*, 16502–16513.

(27) Keutsch, F. N.; Saykally, R. J. Water clusters: Untangling the mysteries of the liquid, one molecule at a time. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 10533–10540.

(28) Zhao, Y.; Schultz, N. E.; Truhlar, D. G. Design of density functionals by combining the method of constraint satisfaction with parametrization for thermochemistry, thermochemical kinetics, and noncovalent interactions. *J. Chem. Theory Comput.* **2006**, *2*, 364–382.

(29) Carey, D. M.; Korenowski, G. M. Measurement of the Raman spectrum of liquid water. *J. Chem. Phys.* **1998**, *108*, 2669–2675.

(30) Flanagin, L. W.; Balbuena, P. B.; Johnston, K. P.; Rossky, P. J. Temperature and density effects on an S_N2 reaction in supercritical water. *J. Phys. Chem.* **1995**, *99*, 5196–5205.

(31) Balbuena, P. B.; Johnston, K. P.; Rossky, P. J. Molecular dynamics simulation of electrolyte solutions in ambient and supercritical water I. Ion solvation. *J. Phys. Chem.* **1996**, *100*, 2706–2715.

(32) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F. AM1: A general purpose quantum mechanical molecular model. *J. Am. Chem. Soc.* **1985**, *107*, 3902–3909.

(33) Hehre, W. J.; Radom, L.; Schleyer, P. v. R.; Pople, J. A. *Ab Initio Molecular Orbital Theory*; John Wiley & Sons: New York, 1986.

JP805499M