The Influence of Environment on Terahertz Spectra of Biological Molecules

Alexei Bykhovski*,†,‡ and Boris Gelmont‡

Department of Electrical & Computer Engineering, North Carolina State University, Raleigh, North Carolina 27695, and Department of Electrical and Computer Engineering, University of Virginia, Charlottesville, Virginia 22904

Received: February 18, 2010; Revised Manuscript Received: June 30, 2010

The variability of molecular vibrations and low terahertz spectra of biological molecules depending on the three-dimensional structure of molecular clusters, chemical bonding, and molecular concentration in the surrounding media is studied using computer simulations. The resonant terahertz spectra of biological molecules and their associations are described within the framework of molecular mechanics using an all-atom molecular mechanical force field for proteins and nucleic acids. Both the absolute values of absorption coefficients and their spectral properties are considered for murein-lipoprotein and thioredoxin of *E. coli* and models of bacterial DNAs using energy minimization and molecular dynamics. The obtained results indicate that structural changes introduced by chemical reactions and molecule associations can strongly affect terahertz spectra, causing significant changes in absorption peak intensities and shifts in peak positions. Terahertz light absorption intensities of studied proteins are predicted to be strongly affected by solvents.

Introduction

Terahertz (THz) spectroscopy of biological (bio) molecules can provide important information on their structure, dynamics, chemical interactions, and identity.[1−8] Experimental techniques were applied to measure THz spectra of diluted solutions, powders, thin films, bacterial spores, and cells. Most of the THz simulations for noncrystals focused on single molecules. Moreover, the emphasis was made on transmission/absorption line positions which allowed many important features of the low THz resonant absorption of biomolecules to be uncovered.[1−7] In particular, molecular dynamical (MD) simulations were performed to study biomolecules in water at room temperature (RT) conditions.[1,3,4] The energy minimization technique was also extensively used for single biomolecule THz research.[5,12,13] Since conformations of biomolecules are sensitive to changes in their environment and biomolecules can readily interact with each other, realistic computer simulations are desirable that go beyond a single molecule approximation. In particular, the absolute value of light absorption is an important gauge of the optical activity of molecules and their state. For understanding underlying biochemical processes and for developing efficient THz detection techniques, it is important to study different sources of variability of THz spectra of biomolecules and to explore specific spectral changes that might be caused by the environment. Spectral signature variability caused by solvents and interactions via intermolecular hydrogen bonds or by subtle changes in chemical composition is the focus of the present research. In this work, both THz spectra and absolute values of absorption coefficients are obtained for bacterial DNA models[5] and *E. coli*'s proteins, thioredoxin and murein-lipoprotein. Thioredoxin, a commonly occurring electron transport protein, was previously studied both experimentally and theoretically in its oxidized form.[1] The scope of the present work includes both the oxidized form thioredoxin−(S_2) and the reduced form thioredoxin−$(SH)_2$. Furthermore, the effect of hydrogen bonds on THz spectra is studied for molecular associations (dimeric and trimeric) formed by the murein-lipoprotein of *E. coli* and for the *E. coli* DNA model. The obtained simulation results suggest that molecule associations have a strong impact on absorption spectra of biomolecules and that they can result in up to 20-fold reduction in absorption peak intensities at THz frequencies. The effect of dielectric properties of organic solvents on terahertz absorption spectra of bacterial proteins is studied using molecular dynamical (MD) simulations and normal-mode analysis with an explicit inclusion of solvent molecules. Three solvents are considered, such as chloroform, methanol, and water. A solute is represented by the thioredoxin and the murein-lipoprotein. Both simulation techniques indicate that a predicted absorption line intensity decreases with a decrease in the dielectric constant of a solvent with the lowest intensities obtained for chloroform and the highest for water.

Absorption Coefficient Simulations

In the previous THz research, the emphasis was made on the absorption peak positions.[12,13] In this work, both the peak positions and the absolute values of the absorption coefficients were obtained as a function of the mass density of a biomolecule for DNA models of *E. coli* and *Bacillus subtilis*, and proteins of *E. coli*. The absorption coefficient, α, is obtained from the imaginary part of the refraction index:[11]

$$\alpha = 4\pi\nu \, \mathrm{Im}(n^*)/c \tag{1}$$

where ν is the frequency of light, c is the speed of light in a vacuum, and $\mathrm{Im}(n^*)$ is the imaginary part of the complex refraction index, n^*:

$$n^* = n(1 + ik) \tag{2}$$

Here, i is imaginary unit and n and k are real functions of a frequency of light.

* To whom correspondence should be addressed. Phone: (919) 593-4400. Fax: (919) 515-5523. E-mail: ab4k@virginia.edu.
† North Carolina State University.
‡ University of Virginia.

10.1021/jp101510k
Published on Web 09/03/2010

Since we directly calculate a dielectric function, rather than the refraction index, it is convenient to express n^* in terms of the complex dielectric function, ε^*:

$$\varepsilon^* = (n^*)^2 \tag{3}$$

For the absorption coefficient, it follows from eqs 1−3

$$\alpha = [2\pi\nu/c]\,\mathrm{Im}(\varepsilon^*)[2^{1/2}(1 - y + (1 + y^2)^{1/2})^{-1/2}]/n^*$$
$$y = \mathrm{Im}(\varepsilon^*)/\mathrm{Re}(\varepsilon^*) \tag{4}$$

Here, $\mathrm{Re}(\varepsilon^*)$ and $\mathrm{Im}(\varepsilon^*)$ are the real and imaginary parts of the dielectric function, correspondingly. In the limit of low absorption, $y \rightarrow 0$:[12]

$$\alpha = [2\pi\nu/c]\,\mathrm{Im}(\varepsilon^*)/n \tag{5}$$

The dielectric function with explicit phonon contributions and phenomenological dissipation factor, γ, can be expressed as follows:[12]

$$\varepsilon^* = \varepsilon_{\mathrm{inf}} + 4\pi N \sum_{\mathrm{K}} (S_k(\omega_k{}^2 - \omega^2 - i\gamma\omega)^{-1}) \tag{6}$$

In eq 6, N is the concentration (number density) of molecules, ω_k is the kth normal mode, $\omega = 2\pi\nu$, S_k is the oscillator strength of the kth normal mode, and $\varepsilon_{\mathrm{inf}}$ is a high-frequency dielectric constant (at frequencies much higher than all ω_k). The oscillator strength is calculated from the normalized amplitude of dipole moment deviation vectors, $\mathbf{P}_k$, for obtained normal modes, $S_k = (\mathbf{P}_k)^2$.[12] If light is polarized along direction $\mathbf{x}$, $S_k = (P_{kx})^2$. As a result of each simulation, THz spectra for three perpendicular polarizations are obtained along with the averaged spectrum that has an oscillator strength $\langle S_k \rangle = [(P_{kx})^2 + (P_{ky})^2 + (P_{kz})^2]/3$.

The line broadening is represented by the $(i\gamma\omega)$ term in eq 6. The damping, γ, depends on a normal mode and it can be represented by a complex function of frequency to account for dispersion and damping of phonons.[11,14,15] However, the calculation of damping and the line shape is beyond the scope of the present work, since it cannot be obtained in the harmonic approximation. Therefore, dissipation has to be determined from the comparison between the theory and experiment. This was done for a number of biological molecules in water.[1,3−5,12,13,16] It was found that, in the first approximation, the damping parameter, γ, can be considered frequency independent and equal to 0.5 cm^{-1} in the low-THz frequency range of interest. The positions of the resonance peaks are not substantially affected by this parameter. It is also anticipated that the changes in the environment considered in this work do not introduce dramatic changes in the damping. Within the framework of the harmonic approximation, any corrections to this assumption can be done by comparing experimental data with the theory.

Notice that the absolute value of the absorption coefficient depends on the concentration of molecules (see eq 6). In the low concentration limit (a diluted solution limit), one can neglect interactions between different molecules and use single molecule results for molecular vibrations to obtain total absorption. However, at higher concentrations or if the equilibrium is shifted toward molecule associations, molecule interactions have to be considered explicitly, since they can alter atomic vibrations and oscillator strength. In this work, molecule interactions between biomolecules are considered for the murein-lipoprotein and a DNA model.

The all-atom Amber molecular mechanical force field (ff03)[10] was used in all structure simulations. The molecular vibrations were simulated using either energy minimizations with the normal-mode analysis or molecular dynamics (MD) with the quasi-harmonic analysis. The quasi-harmonic approximation is used to extract spectra of biomolecules in bulk solutions at atmospheric pressure and room temperature conditions from generated MD atomic trajectories.[1,3,4] The normal-mode analysis cannot be applied here because of diffusion of solvent molecules, since its validity is restricted to an analysis of small atomic displacements from an equilibrium atomic configuration that corresponds to a potential energy minimum. On the other hand, the normal-mode analysis is utilized to generate THz spectra for partially solvated (dry) biomolecules. In this case, the effect of temperature on spectra was not considered. Each bacterial DNA structure was obtained as in our previous work[5] from the energy minimizations of the 20 base pair model structures with base pair sequences that were determined using the entire bacterial DNA sequence information to approximate DNA composition and base correlations. Specifically, for *E. coli* and *Bacillus subtilis*, the sequences AAGTCCAGCTTTCAGACTGG and CATGTCTGAAAGCAGTACTT were used, respectively.[5] The protein thioredoxin was modeled in TIP3P water using the molecular dynamical (MD) simulations with realistic environmental conditions (at atmospheric pressure and room temperature), and its atomic vibrations were obtained in the quasi-harmonic approximation from the generated trajectories.[1] The dielectric function and absorption coefficient for different mass densities were obtained in the model of hard (nonpolarizable) ions, so that $\varepsilon_{\mathrm{inf}} = 1$ and each atom was assigned a fixed effective charge. In Figure 1a, the predicted absorption coefficient's spectrum of the thioredoxin is shown below 1 THz for 1, 10, and 100 g/L. In Figure 1b, the absorption coefficient of the thioredoxin at two peaks, 15.66 and 22.5 cm^{-1}, is plotted for a range of densities from 1 to 400 g/L. The dependence of the absorption coefficient on density exhibits a frequency-dependent sublinear character. In Figures 2 and 3, similar plots are presented for the DNA models of *E. coli* and *Bacillus subtilis*, respectively.

Molecule Associations

Depending on the environment, biological molecules form hydrogen-bonded molecule associations. Their impact on THz spectra was studied using computer simulations for murein-lipoproteins originating from *E. coli* and for the *E. coli* DNA model. The murein-lipoprotein of *E. coli* is one of the most abundant proteins in *E. coli* bacteria with 7.2×10^5 molecules per cell that are localized in the outer membrane.[17] It exists in two forms, the bound form and the free form. Roughly twice as much of the free form is found as of the bound form. The bound form of the murein-lipoprotein makes up about 10% of the cell wall in *E. coli*.[18] The murein-lipoprotein can exist in monomeric, dimeric, and trimeric structures. It has a helical symmetry and no electric charge. The trimer creates a cylinder 83 Å in length with a circular cross section of 21 ± 24 Å. The refined atomic coordinates from X-ray data[13] were used as input to the THz spectra simulations. The stable conformations of the murein-liporpotein structures have been obtained by minimizing the energy to the root mean square (rms) of the Cartesian elements of the force below 10^{-5} kcal/mol/Å. The predicted effective potential energy of the trimer (i.e., the energy per molecule unit in the trimer) is lower than the monomer's energy

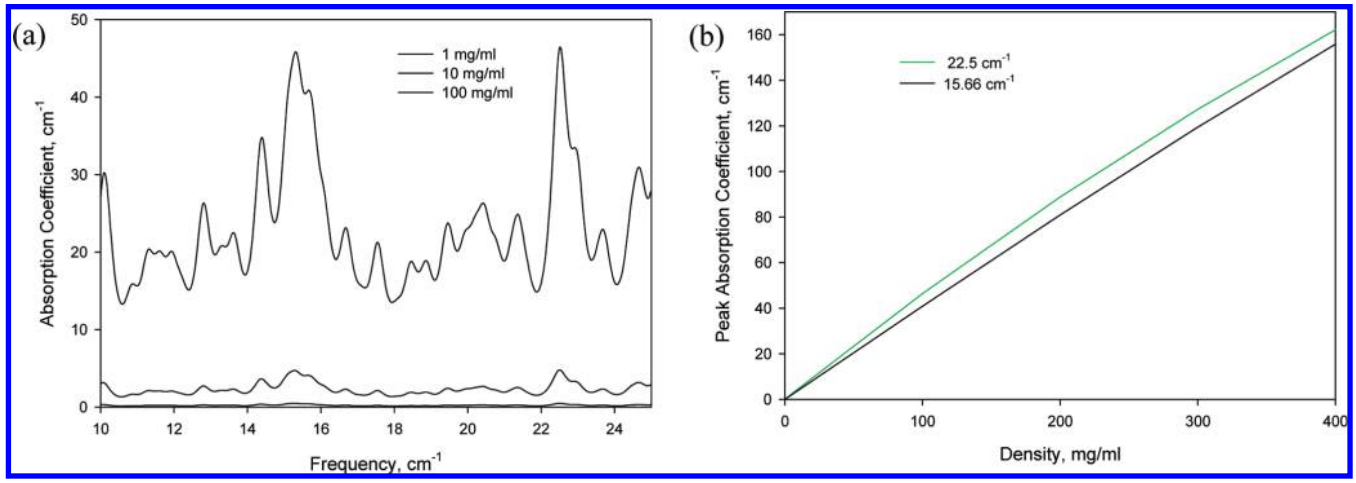

Figure 1. Oxidized thioredoxin from *E. coli*. Molecular dynamical simulations with TIP3P explicit water. Absorption coefficients averaged over light polarizations. Absorption coefficient spectrum (a); peak absorption coefficients vs density (b).

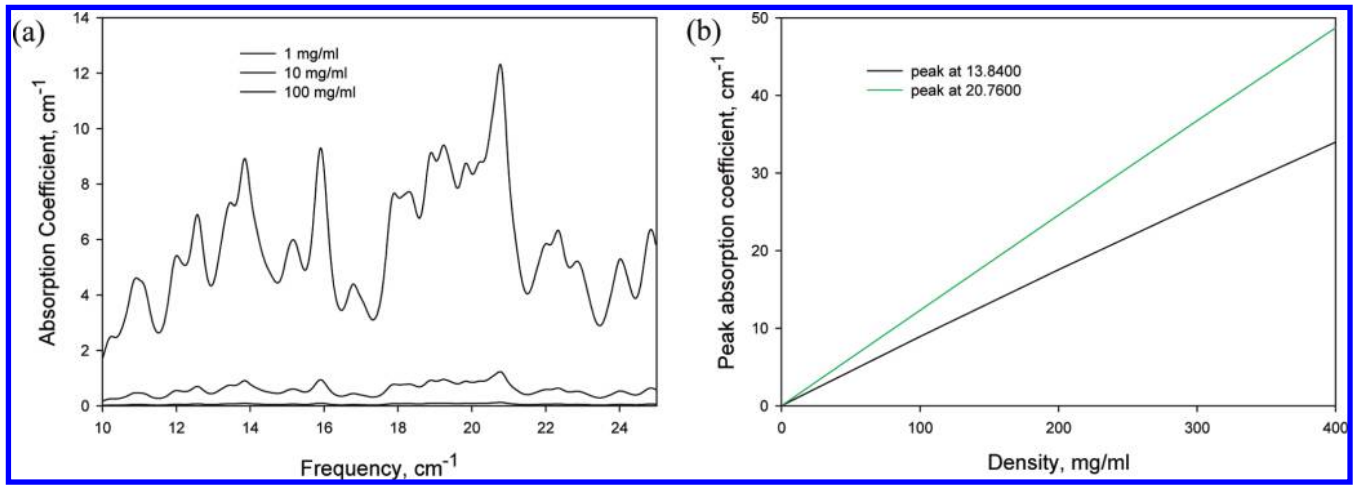

Figure 2. *E. coli* DNA model. Amber 8 energy minimization results. Absorption coefficients averaged over light polarizations. Absorption coefficient spectrum (a); peak absorption coefficients vs density (b).

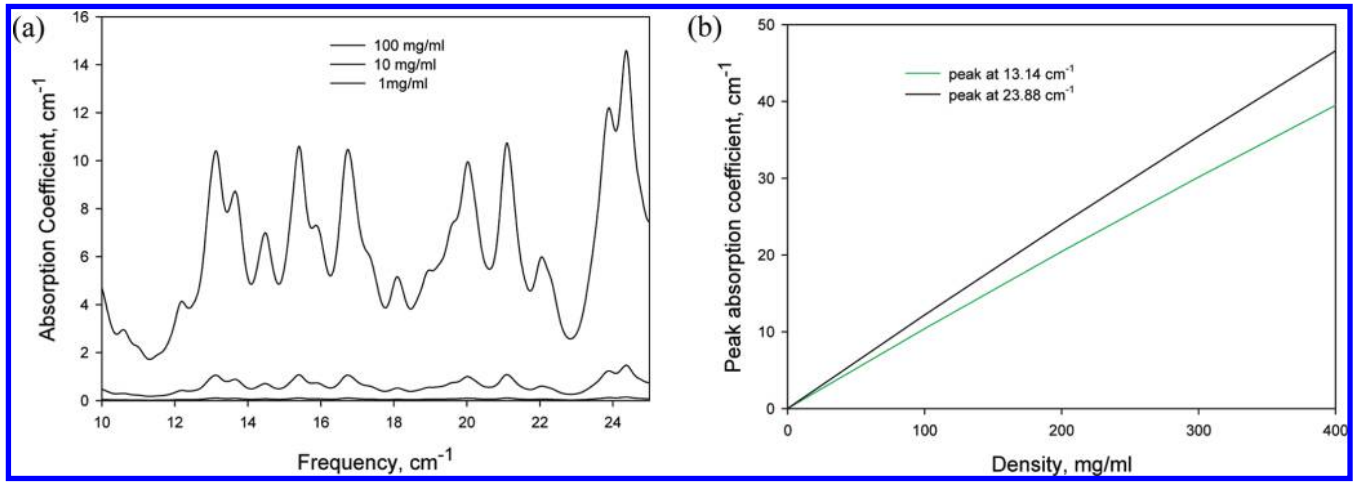

Figure 3. *Bacillus subtilis* DNA model. Amber 8 energy minimization results. Absorption coefficients averaged over light polarizations. Absorption coefficient spectrum (a); peak absorption coefficients vs density (b).

by approximately 11 eV if a zero point vibrational energy is not counted. Including the zero point vibrational energy and the thermal vibrational energy at room temperature (RT), the per-molecule energy of the trimer is lower than the monomer's energy by 10.6 eV. Therefore, the simulations predict that the trimer is stable at RT, in accordance with the available experimental data.[17] In Figure 4, the resulting structures of the murein-lipoprotein trimer and monomer are plotted. In Figure 5, a view at the two ends of the trimer is provided with an initial trimeric structure (thin lines) and a final optimized structure being superimposed. The low THz absorption coefficients were calculated in a broad mass density range (1−400 g/L). In Figure 6a, the absorption coefficients of the trimer (black lines) and the monomer (green lines) averaged over orientations are shown below 1 THz. There is a dramatic difference in absorption strength and also in line positions between absorption coef-

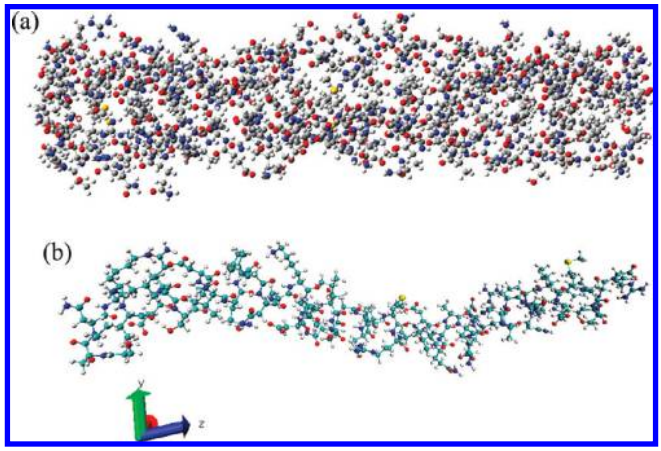

Figure 4. Optimized murein-lipoprotein (a) trimer and (b) monomer from *E. coli*.

ficients of the trimer and the monomer, with the trimer's absorption being almost everywhere much lower, with only a few exceptions such as the peaks at 17.3 and 2.2 cm^{-1}. For example, a peak at 13 cm^{-1} is 17 times weaker for the trimer. Also, monomer lines are much stronger at 5, 11, 13, 15, 19, 21, and 23.7 cm^{-1}. In Figure 6b−d, the polarized absorption results are shown for *X*, *Z*, and *Y*, correspondingly, along with the average absorption. Results for different polarizations also show a contrast between the monomer's and trimer's sub-THz signatures.

The predicted dimer potential energy per unit molecule is lower than the monomer energy by 9%; however, the trimer energy per unit molecule is even lower. In Figure 7, the averaged over orientations absorption coefficient spectra of the dimer (blue lines) and the monomer (green lines) are shown below 1 THz for 10, 100, and 200 g/L. Notice that most of the dimer's peaks are much weaker for the same mass density than their monomer's counterparts. For example, the characteristic peaks at 11, 13.1, and 21.6 cm^{-1} are 5−6 times weaker. At the same time, the dimer has spectral features at 12.1, 18.4, and 25 cm^{-1} that are not predicted for the monomer. In Figure 8, *X*-polarized and averaged absorption spectra are plotted for the monomer (red), the dimer (blue), and the trimer (black). The predicted monomer's absorption intensities are much higher in the sub-THz regime than those of both molecular associations with the exception of the absorption peaks of the dimer at 18.4 and 25 cm^{-1}.

The influence of water on THz spectra of murein-lipoprotein associations was also studied using energy minimizations. The 61 water molecules (waters) were added to each strand's surface. The resulting THz spectra of the trimer and the monomer averaged over light polarizations for the mass density 1 g/L are plotted in Figure 9. The predicted spectra show much stronger absorption peaks for the monomer for most of the frequencies of the selected 0−90 cm^{-1} frequency range. A few exceptions, in particular near 41 and 65 cm^{-1}, do exist where trimer absorption is about as strong or somewhat stronger than the monomer's. Therefore, both dry and wet murein-lipoprotein associations tend to reduce THz absorption intensities.

Hydrogen-bonded molecule associations for the 20 base pair DNA model of *E. coli* that was described in the previous section also demonstrated a strong influence on THz absorption spectra. In Figure 10, the single stranded *E. coli* DNA spectrum that was averaged over light polarizations is plotted vs the spectrum for the double stranded *E. coli* DNA for the mass density 1 g/L. In contrast to the obtained murein-lipoprotein simulation results, the modeled hydrogen-bonded association of DNA strands (blue lines in Figure 10) does not uniformly weaken the THz absorption in the considered 0−3 THz range. Moreover, the double stranded DNA is predicted to have 4−5 times stronger absorption peaks between 74 and 95 cm^{-1}. These differences in absorption, in particular the differences between murein-lipoprotein and DNA, might be caused by the symmetry of the associations. The higher symmetry structure, the trimer, produces the weakest oscillator strengths and is predicted to have approximately 3 times weaker absorption in the 0−3 THz range than the double stranded DNA for the same material density. In other words, because it registers collective motions, THz spectroscopy is sensitive to the symmetry of the entire biostructures.

It follows from the obtained results that both protein-based and nucleic-acid-based molecular associations can strongly influence THz absorption characteristics. Therefore, any process that leads to a structural instability and strand separations such as denaturing which can be caused by temperature or other factors can be detected by THz spectroscopy methods. The issue of thermal stability of molecular associations in water was explored using MD simulations. The thermal stability of the murein-lipoprotein trimer was studied using particle mesh Ewald molecular dynamics (PMEMD).[10] The results of a 15.3 ns run in a periodic box of 2326 TIP3P waters with 2 fs step with the SHAKE algorithm at RT and atmospheric pressure conditions confirmed its stability. No strand separation was obtained. The calculated root-mean-square deviation (rmsd) for the backbone atoms shows an average of 1.45 Å and a standard deviation of 0.44 Å. Our preliminary simulation results for elevated tem-

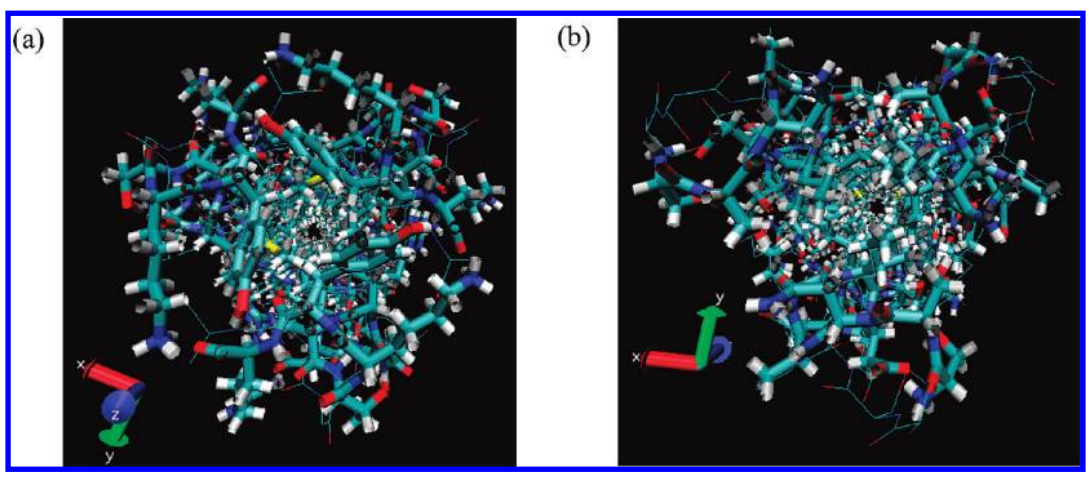

Figure 5. Murein-lipoprotein trimer from *E. coli*. View along the trimer axis. Energy minimization results (bonds); initial structure (thin lines).

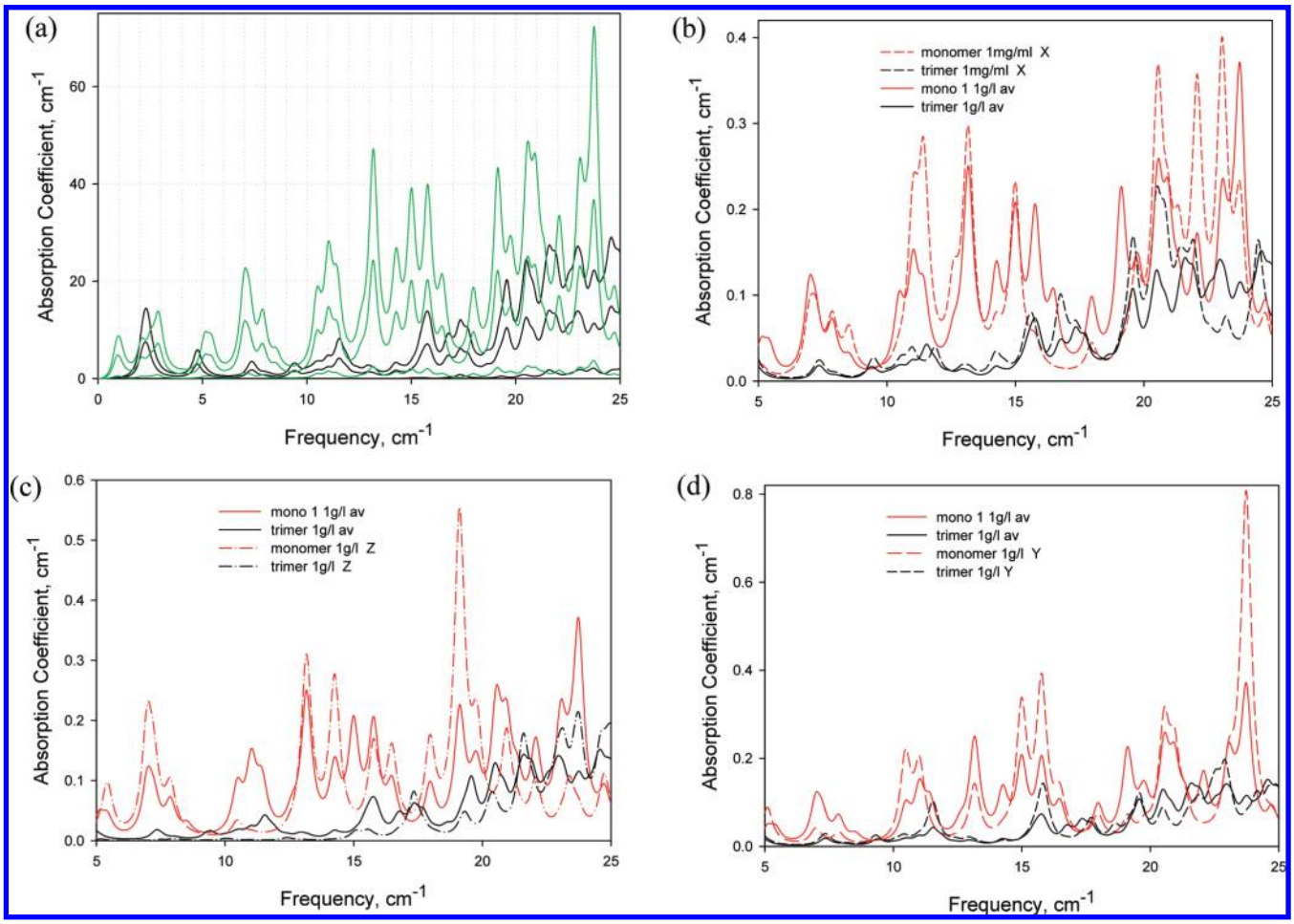

Figure 6. Murein-lipoprotein from *E. coli*. Amber 9 energy minimization results. (a) Absorption coefficient for 10, 100, and 200 g/L. Green lines - monomer, black lines - trimer. (b−d) Absorption spectrum averaged over light polarizations (solid lines) and (b) *X*-polarized (dashed lines), (c) *Y*- polarized, and (d) *Z*-polarized. Red lines - monomer, black lines - trimer.

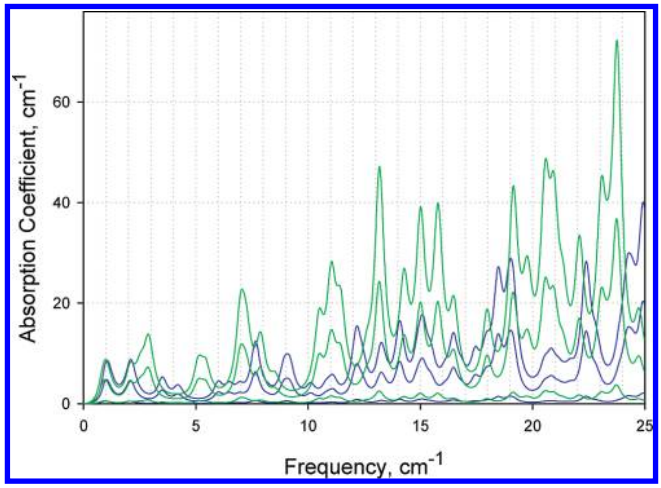

Figure 7. Murein-lipoprotein from *E. coli*. Calculated absorption coefficient for 10, 100, and 200 g/L; green - monomer, blue - dimer.

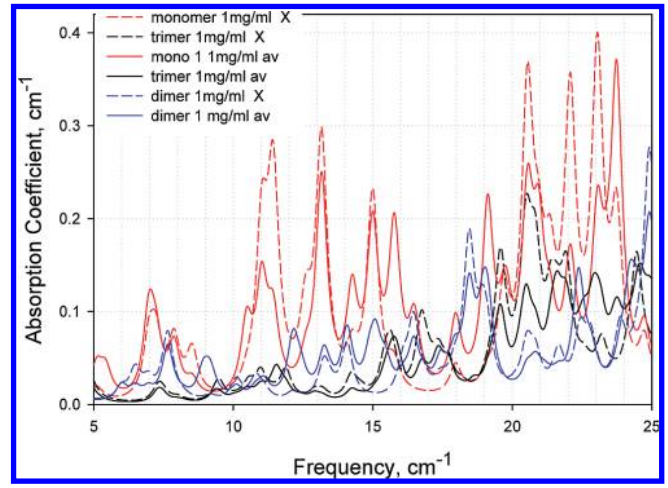

Figure 8. Murein-lipoprotein from *E. coli*. Predicted absorption spectra averaged over light polarizations (solid lines) and *X*-polarized (dashed lines). Red - monomer, blue - dimer, black - trimer.

peratures indicate the trimer's stability at least to 330 K in pure water. The predicted thermal stability of the trimer at RT and elevated temperature in water is in agreement with the experimental data.[17]

The Effect of Organic Solvents

A natural environment of biomolecules consists of water-based solvents (inside cells) or protein-based networks (in spores and cell walls) that have different dielectric properties. In order to describe these complex environments, computer simulations have to go beyond a single molecule approximation. Recently, the properties of several biomolecules in water at sub-THz frequencies and RT have been investigated and as a result both experimental and theoretical THz spectra of oxidized *E. coli*'s thioredoxin and the $tRNA_{tyr}$ specific for tyrosine as well as an artificial DNA decamer duplex have been obtained.[1,3,4] In these

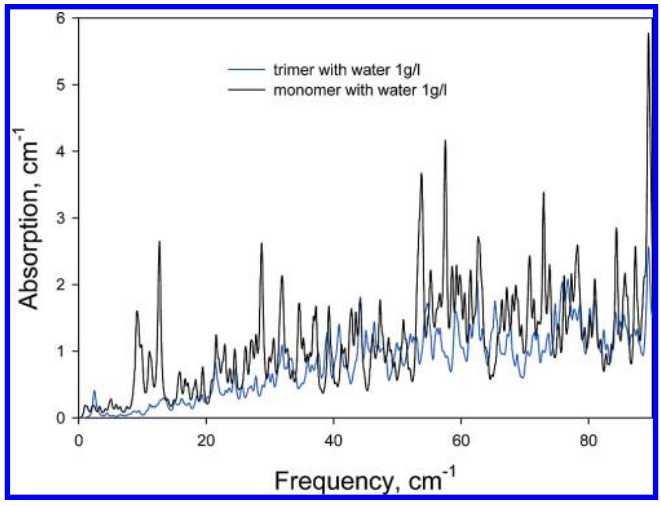

Figure 9. Murein-lipoprotein trimer from *E. coli*. Blue line - trimer with 183 water molecules added. Black line - monomer with 61 waters added. Absorption spectrum averaged over light polarizations. The mass density is 1 g/L.

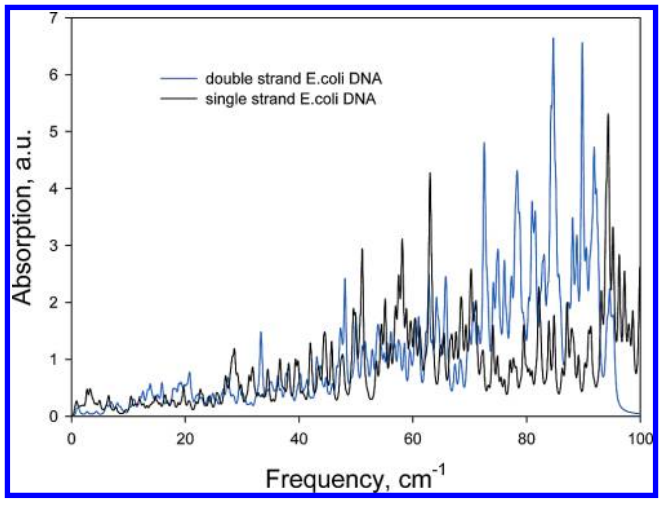

Figure 10. Single stranded *E. coli* DNA (black line) vs double stranded *E. coli* DNA (blue line). The mass density is 1 g/L. The absorption spectrum is averaged over light polarizations.

simulations, water molecules were introduced explicitly via the TIP3P model and a strong dependence of biomolecule properties on water was demonstrated. In the present research, the effect of organic solvents with different dielectric properties on picosecond dynamics and THz absorption characteristics of *E. coli*'s thioredoxin is explored using two types of simulations, MD that is followed by a quasi-harmonic analysis and an energy minimization with a subsequent harmonic analysis. The latter technique was also applied in the case of the murein-lipoprotein. All MD simulations have followed a simulation protocol that was previously developed to obtain THz spectra of the oxidized thioredoxin in water that were in good agreement with the experimental data.[1] In particular, boxes for periodic simulations were generated with one thioredoxin and 1455 methanol molecules or 689 chloroform molecules. Preliminary minimizations were followed by heating to RT and by equilibrations with subsequent production runs at constant temperature and pressure conditions. In Figure 11, the predicted absorption spectra averaged over orientations for thioredoxin/chloroform (dark green), thioredoxin/methanol (blue), and thioredoxin/water (red) systems are plotted in the 0−3 THz (0−100 cm^{-1}) range. The predicted absorption of thioredoxin in water (dielectric constant 78) is the highest overall, while thioredoxin in chloroform

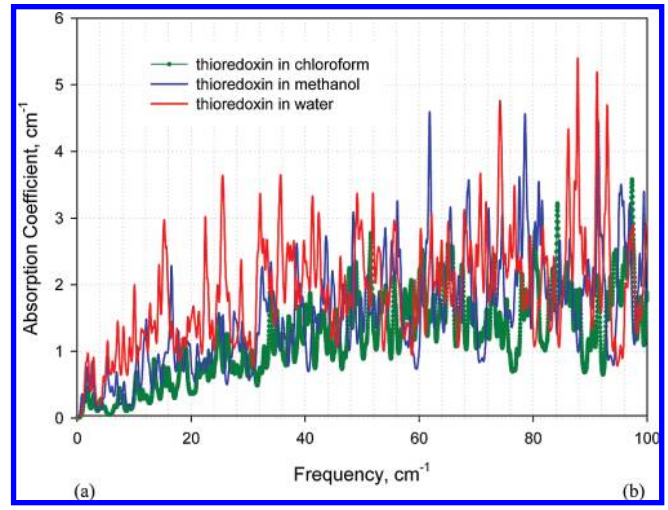

Figure 11. Predicted absorption spectra averaged over orientations for thioredoxin/chloroform (dark green), thioredoxin/methanol (blue), and thioredoxin/water (red). MD simulations, 300 K, 1 atm.

(dielectric constant 4.8) shows the weakest overall absorption. The thioredoxin in methanol (dielectric constant 32.7) exhibits intermediate absorption intensities with the exception of strong lines around 62 and 78, 95, and 99 cm^{-1}. This apparent trend in the protein absorption might be related to stronger disturbances in the protein structure that lead to higher effective dipole moments caused by more polar solvents. In order to explore this issue further and to study the sensitivity of THz spectra to organic solvents, THz absorption spectra of only partially solvated biomolecules were simulated as well. One of the questions here is how distinguished are the THz absorption characteristics of almost dry proteins and fully solvated ones. Therefore, the energy minimizations of the thioredoxin and the murein-lipoprotein with thin shells formed by the same organic solvents were performed. In Figure 12, final optimized structures of thioredoxin/chloroform (a), thioredoxin/methanol (b), and thioredoxin/water systems (c) with solvent shells (colored spheres) are shown. Corresponding terahertz absorption spectra for 0−2 THz that were obtained from the normal-mode analyses are plotted in Figures 13 and 14. The predicted absorption spectra averaged over orientations for thioredoxin/chloroform (dark green), thioredoxin/methanol (black), and thioredoxin/water (red) show the same trend as the RT MD-based spectra of Figure 11. Therefore, this absorption reduction in the low dielectric solvent is not linked to thermal effects. In Figure 13, the total absorption of the solute/solvent system is shown. The thioredoxin/methanol is predicted to have almost as strong absorption as the thioredoxin/water system with some of the peaks (around 54, 56, 60, and 62 cm^{-1}) stronger than those for thioredoxin/water. The thioredoxin/chloroform absorption intensity lines (green) are roughly 3 times weaker than those for more polar solvents. The thioredoxin-only contributions into the light absorption for the same systems are plotted in Figure 14 which demonstrates the same qualitative trend in absorption intensities. However, in Figure 14, peaks for the thioredoxin from the thioredoxin/water system are noticeably stronger than those for the thioredoxin from the thioredoxin/methanol system. If solvents did not contribute to absorption at these frequencies, Figures 13 and 14 would be identical. The differences in intensities between Figures 13 and 14 point to some solvent shell contributions. Finally, similar results were obtained for *E. coli*'s murein-lipoprotein with organic solvent shells. In Figure 15, the 0−2 THz absorption coefficient spectra of the murein-lipoprotein trimer are plotted. The absorption line

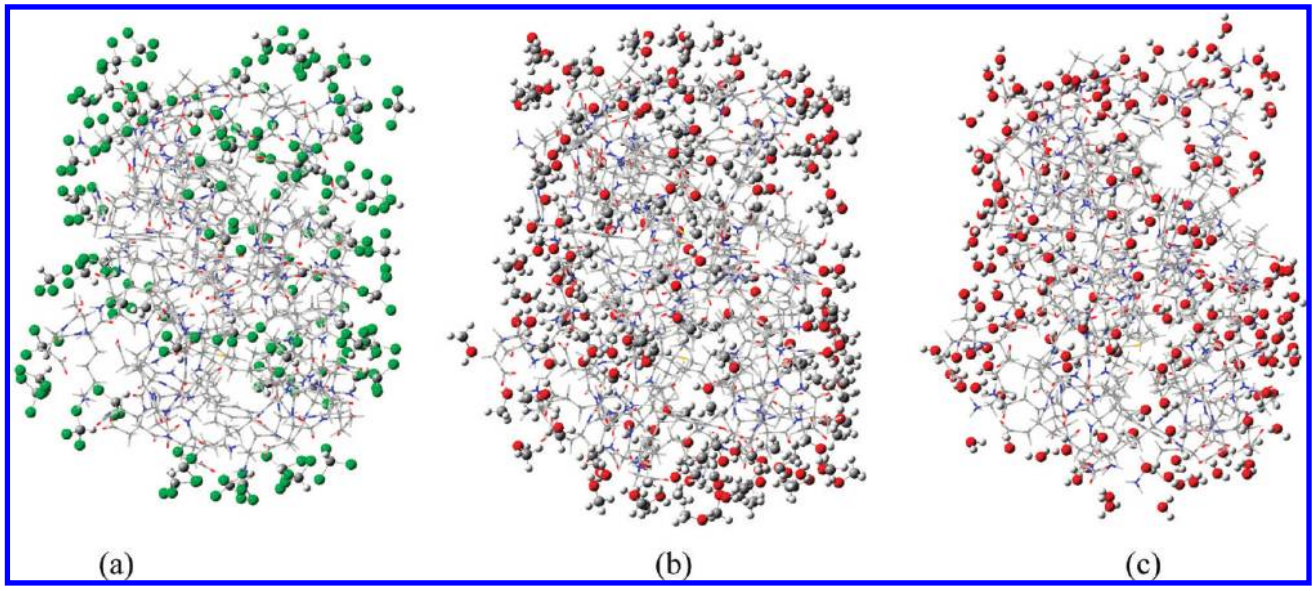

Figure 12. Final optimized structures of thioredoxin/chloroform (a), thioredoxin/methanol (b), and thioredoxin/water systems (c) with limited amounts of solvents (colored spheres). Chlorine - green, oxygen - red, carbon - gray, hydrogen - white.

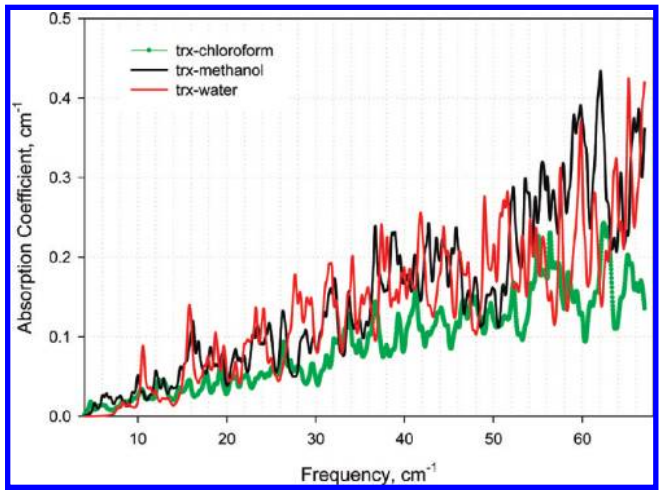

Figure 13. The predicted absorption spectra averaged over orientations for thioredoxin/chloroform (dark green), thioredoxin/methanol (black), and thioredoxin/water (red). Based on energy minimization with normal-mode analysis.

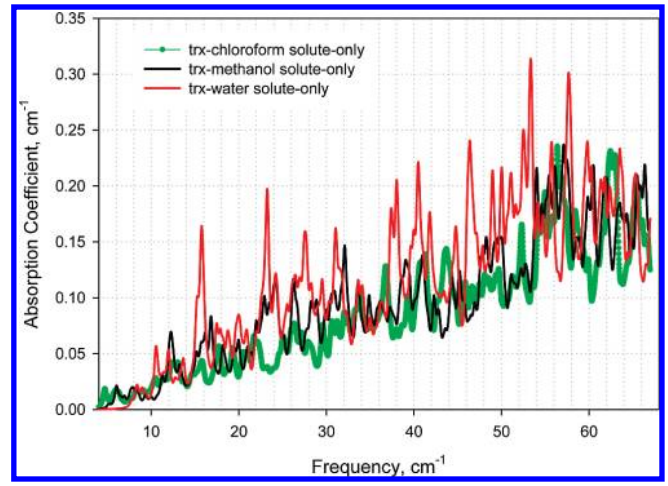

Figure 14. Thioredoxin-only absorption spectra averaged over orientations for thioredoxin/chloroform (dark green), thioredoxin/methanol (black), and thioredoxin/water (red) systems.

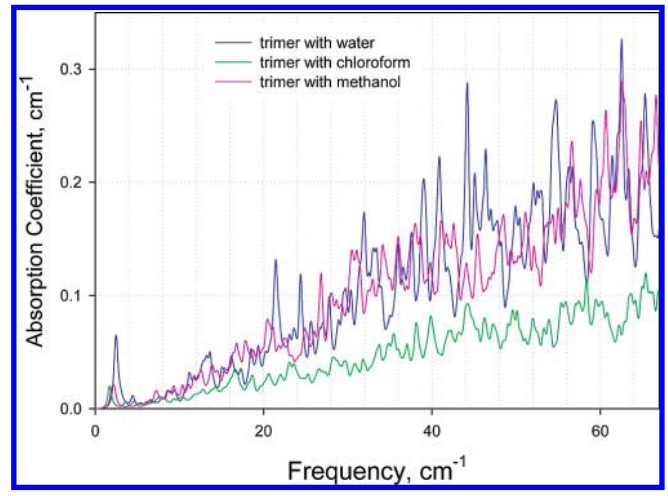

Figure 15. Predicted absorption spectra for murein-lipoprotein trimer with water, methanol, and chloroform shells.

intensities for the trimer/chloroform system are 2−3 times lower than for the trimer/water system.

The obtained changes in absorption intensities of the thioredoxin and murein-lipoprotein with a change in a solvent's dielectric constant may provide clues to understanding absorption in biological matter. In particular, the effective dielectric constants in low-dielectric protein environments,[19] such as dry bacterial spores (∼4), are close to the dielectric constant of chloroform. On the other hand, hydrated more polar environments typical for vegetative cells can have much higher dielectric constants. Assuming the predicted absorption trend holds for broader classes of proteins, one could expect a THz absorption decrease with a decrease in the water content of a biomaterial, since the overall polarity of the system would be reduced. Notice, however, that the same damping coefficient was used in all presented simulations. Since changes in damping might influence the absorption line intensities, combined experimental and theoretical studies are required to address the issue of damping in different environments.

Reduced Thioredoxin of *E. coli*

There are two equally important forms of the thioredoxin, oxidized and reduced.[9] Reduced *E. coli* thioredoxin has two

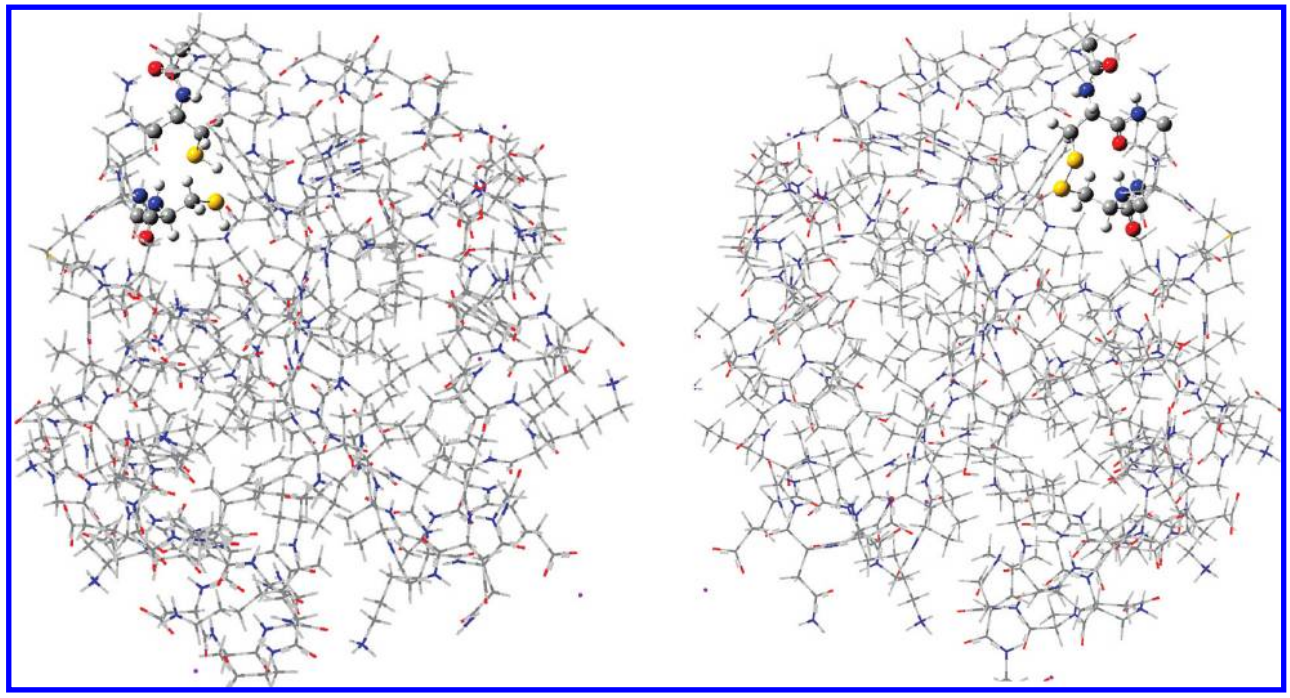

Figure 16. Optimized structures of the thioredoxins from *E. coli*. Reduced (left), oxidized (right). Spheres represent atoms at key sulfur sites: S−H, S−H (left), and S−S (right). Sulfur - yellow, carbon - gray, oxygen - red, hydrogen - white, nitrogen - blue.

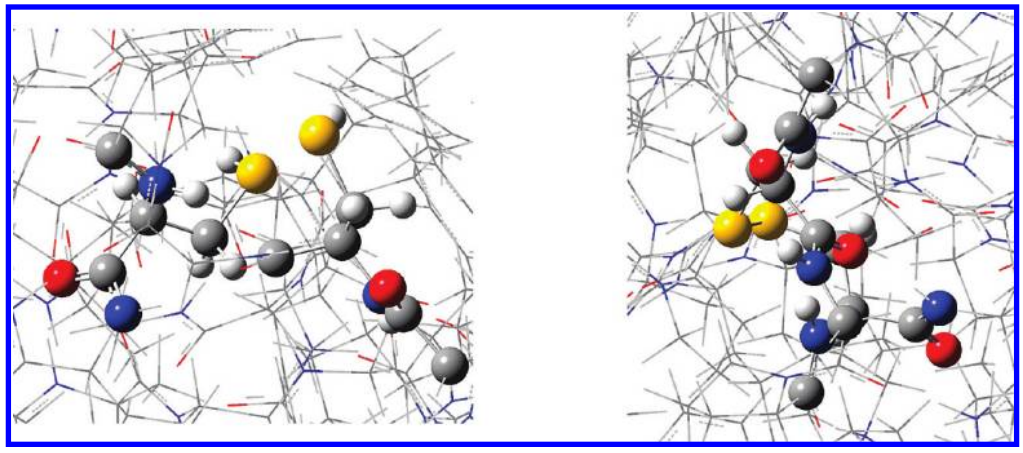

Figure 17. Optimized structures of the thioredoxins from *E. coli*. Structure fragments. Reduced (left), oxidized (right). Spheres represent atoms at key sulfur sites. S−H, S−H (left), and S−S (right).

sulfhydryl groups (−SH) which are oxidized to a disulfide bond (−S−S−), forming the oxidized thioredoxin. In *E. coli* and in mammals, this biochemical process is used to form deoxyribonucleotides for DNA synthesis from the ribonucleotides. In the present study, a dry thioredoxin fully neutralized by four sodium ions was considered. The energy minimization procedure was utilized to find its equilibrium structure. The resulting optimized structures of reduced and oxidized thioredoxins are shown in Figures 16 and 17. The differences in conformations at the sulfur site can clearly be seen. The normal-mode analyses on optimized structures were performed, and the absorption spectra were calculated within 0−3 THz. The results for the averaged over orientations absorption intensities are plotted in Figure 18 for the reduced (blue) and oxidized (red) forms. The predicted spectra of the reduced and oxidized thioredoxins are distinguishable in the entire 0−3 THz range, especially between 16−20, 24−28, 40−44, at 60, and 76−100 cm^{-1}. In Figure 19, the predicted spectra for the 10−25 cm^{-1} sub-THz frequencies are plotted. There is no peak at 11 cm^{-1} in the reduced form. The strong peaks at 16.8 and 18.3 cm^{-1} exist in the predicted reduced

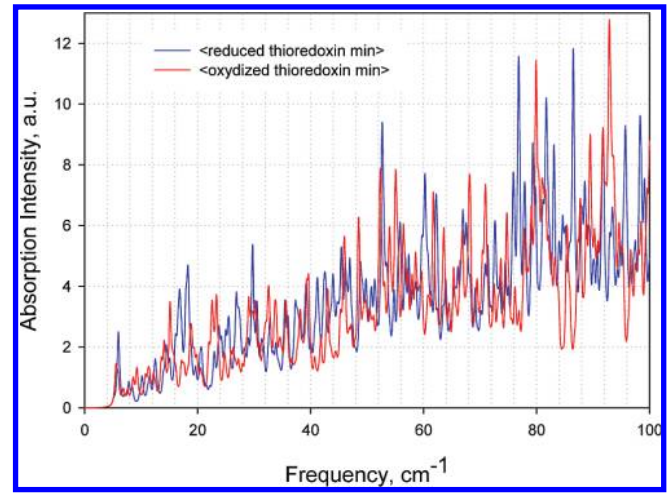

Figure 18. The average absorption intensity of the energy-minimized dry thioredoxin from *E. coli*. Reduced thioredoxin (blue), oxidized (red).

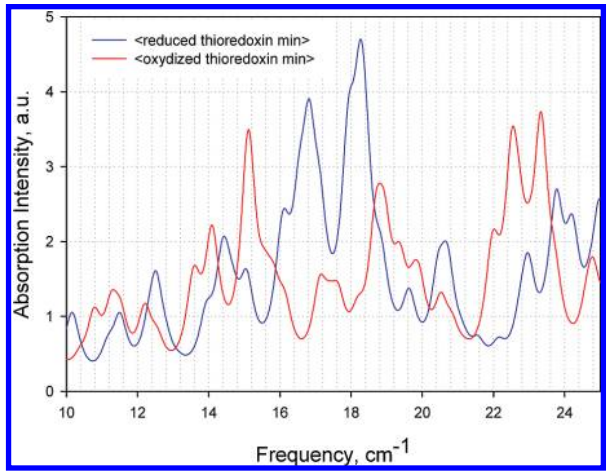

Figure 19. Average absorption intensity of the energy-minimized dry thioredoxin from *E. coli*. Reduced thioredoxin (blue), oxidized (red). Sub-THz part of Figure 18.

thioredoxin spectrum. The somewhat similar double feature of the oxidized form is shifted to higher frequencies by ~0.5 cm^{-1} and is weaker. The dominant peak of the oxidized form at 15.2 becomes a secondary feature in the reduced form. The double peak of the oxidized form at 22−23 cm^{-1} is shifted to higher frequencies by 0.4−0.5 cm^{-1}, and its signature is changed. Therefore, the THz spectra of the thioredoxin are predicted to be extremely sensitive to the conformational changes surrounding the double sulfur site.

Conclusions

In this work, the influence of the environment on THz spectra of several nucleic acids and proteins was considered using molecular mechanical simulation techniques within the framework of harmonic and quasi-harmonic approximations. The absorption coefficients for DNAs of *E. coli* and *Bacillus subtilis* and for proteins of *E. coli* are obtained. The absorption spectra show a frequency-dependent sublinear dependence on molecular mass densities for low densities of biomaterials. THz spectra of the thioredoxin are predicted to be extremely sensitive to the conformational changes surrounding the double sulfur site. The effect of weak hydrogen bonds on THz spectra is studied for molecular associations formed by the murein-lipoprotein and DNA of *E. coli*. The formation of molecule associations of nucleic acids and proteins is predicted to have a strong impact on THz absorption spectra of biomolecules, leading to substantial changes in absorption peak intensities and shifts in line positions. The predicted thermal stability of the studied molecule associations is consistent with the available data. In particular, the murein-lipoprotein trimer is predicted to be stable to at least 330 K. According to the obtained results, processes leading to strand separations such as denaturing can be monitored by THz spectroscopy techniques. The influence of bulk and thin shells of organic solvents on THz spectra of the murein-lipoprotein and thioredoxin is considered using explicit solvent models. The obtained results point to a reduction of THz absorption intensities of studied proteins in low dielectric environments assuming the same damping. Combined experimental and theoretical studies are required to address the issue of damping in different environments.

Acknowledgment. This work is supported by the DTRA CB Basic Research Program, Contract No. HDTRA1-08-1-0038, and by the National Research Council.

References and Notes

(1) Bykhovski, A.; Globus, T.; Khromova, T.; Gelmont, B.; Woolard., D. Resonant Terahertz Spectroscopy of Bacterial Thioredoxin in Water: Simulation and Experiment. *Int. J. High Speed Electron. Syst.* **2008**, *18* (1), 109–117.

(2) Globus, T.; Woolard, D.; Crowe, T. W.; Khromova, T.; Gelmont, B.; Hesler, J. Terahertz Fourier transform characterization of biological materials in a liquid phase. *J. Phys. D: Appl. Phys.* **2006**, *39* (15), 3405–3413.

(3) Li, X.; Globus, T.; Gelmont, B.; Salay, L. C.; Bykhovski, A. Terahertz Absorption of DNA Decamer Duplex. *J. Phys. Chem. A* **2008**, *112* (47), 12090–12096.

(4) Bykhovski, A.; Globus, T.; Khromova, T.; Gelmont, B.; Woolard., D. An Analysis of the THz Frequency Signatures in the Cellular Components of Biological Agents. *Int. J. High Speed Electron. Syst.* **2007**, *17* (2), 225–237.

(5) Bykhovski, A.; Li, X.; Khromova, T.; Globus, T.; Gelmont, B.; Woolard, D.; Samuels, A. C.; Jensen, J. O. THz absorption signature detection of genetic material of *E. coli* and *B. subtilis*. *Chemical and biological standoff detection III*; Boston, MA, October 2005, *5995*, 59950N-1:10.

(6) Globus, T.; Norton, M. L.; Lvovska, M. I.; Gregg, D. A.; Khromova, T. B.; Gelmont, B. Reliability Analysis of THz Characterization of Modified and Unmodified Vector Sequences. IEEE Sens. J., in press.

(7) Korter, T. M.; Plusquellic, D. F. Continuous-wave terahertz spectroscopy of biotin: vibrational anharmonicity in the far-infrared. *Chem. Phys. Lett.* **2004**, *385*, 45–51.

(8) Siegrist, K.; Bucher, C. R.; Mandelbaum, I.; HightWalker, A. R.; Balu, R.; Gregurick, S. K.; Plusquellic, D. F. High-Resolution Terahertz Spectroscopy of Crystalline Trialanine: Extreme Sensitivity to β-Sheet Structure and Cocrystallized Water. *J. Am. Chem. Soc.* **2006**, *128* (17), 5765.

(9) Katti, S. K.; LeMasterf, D. M.; Eklund, H. Crystal Structure of Thioredoxin from Escherichia coli at l.68 A Resolution. *J. Mol. Biol.* **1990**, *212*, 167–184.

(10) Case, D. A.; Pearlman, D. A.; Caldwell, J. W.; Cheatham, T. E., III; Wang, J.; Ross, W. S.; Simmerling, C. L.; Darden, T. A.; Merz, K. M.; Stanton, R. V.; Cheng, A. L.; Vincent, J. J.; Crowley, M.; Tsui, V.; Gohlke, H.; Radmer, R. J.; Duan, Y.; Pitera, J.; Massova, I.; Seibel, G. L.; Singh, U. C.; Weiner, P. K.; Kollman, P. A. *AMBER 8*; University of California: San Francisco, CA, 2004. http://ambermd.org/.

(11) Born, M.; Huang, K. *Dynamical Theory of Crystal Lattices.*; Oxford: New York, 1954.

(12) Bykhovskaia, M.; Gelmont, B.; Globus, T.; Woolard, D. L.; Samuels, A. C.; Duong, T. H.; Zakrzewska, K. Prediction of DNA far-IR absorption spectra based on normal mode analysis. *Theor. Chem. Acc.* **2001**, *106*, 22–27.

(13) Globus, T.; Woolard, D.; Bykhovskaia, M.; Gelmont, B.; Werbos, L.; Samuels, A. THz-frequency spectroscopic sensing of DNA and related biological materials. *Int. J. High Speed Electron. Syst.* **2003**, *13* (4), 903–936.

(14) Ipatova, I. P.; Maradudin, A. A.; Wallis, R. F. Temperature Dependence of the Width of the Fundamental Lattice-Vibration Absorption Peak in Ionic Crystals. II. Approximate Numerical Results. *Phys. Rev.* **1967**, *155*, 882–895.

(15) Monga, M. R.; Jindal, V. K.; Pathak, K. N. Self-Energy of Phonons in an Anharmonic Crystal of Order λ^4. III. Approximate numerical Results for Ionic Crystals. *Phys. Rev. B* **1979**, *19*, 1230–1242.

(16) Xiaowei, L.; Bykhovski, A.; Gelmont, B.; Globus, T.; Woolard, D.; Bykhovskaia, M. Computational methods for analysis of terahertz spectral signatures of nucleic acids. Nanotechnology, 2005. 5th IEEE Conference on; Nagoya, Japan, July 11−15, 2005; pp 375−378. Digital Object Identifier: 10.1109/NANO.2005.1500734.

(17) Shu, W.; Liu, J.; Ji, H.; Lu, M. Core structure of the outer membrane lipoprotein from Escherichia coli at 1.9 å resolution. *J. Mol. Biol.* **2000**, *299*, 1101–1112.

(18) Bosch, V.; Braun, V. Distribution of' mureinlipoprotein between the cytoplasmic and outer membrane of Escherichia coli. *FEBS Lett.* **1973**, *34*, 307–310.

(19) Myers, J.; Grothaus, G.; Narayanan, S.; Onufriev, A. A Simple Clustering Algorithm Can Be Accurate Enough for Use in Calculations of pKs in Macromolecules. *Proteins: Struct., Funct., Bioinf.* **2006**, *63*, 928–938.

JP101510K