

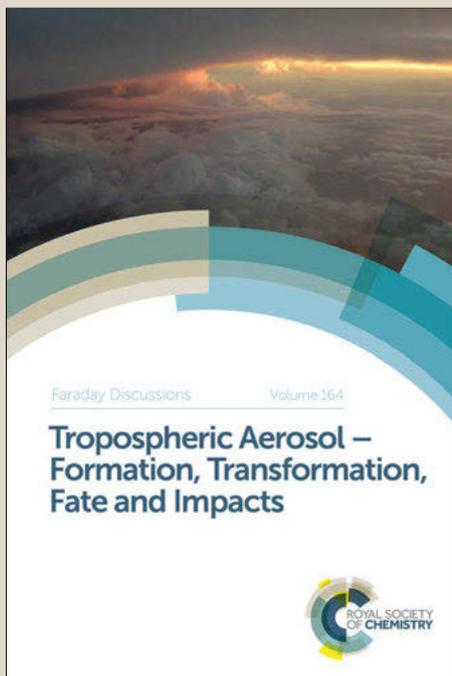
# Faraday Discussions

Accepted Manuscript



This manuscript will be presented and discussed at a forthcoming Faraday Discussion meeting. All delegates can contribute to the discussion which will be included in the final volume.

**Register now to attend!** Full details of all upcoming meetings: <http://rsc.li/fd-upcoming-meetings>



This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

*Accepted Manuscripts* are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.

# Sub-THz specific relaxation times of hydrogen bond oscillations in *E.coli* thioredoxin. Molecular dynamics and statistical analysis.

Tatiana Globus<sup>a,b\*</sup>, Igor Sizov<sup>a,b</sup> and Boris Gelont<sup>a,b</sup>

DOI: 10.1039/b000000x [DO NOT ALTER/DELETE THIS TEXT]

Hydrogen bonds (H-bonds) in biological macromolecules are important for molecular structure and functions. Since interactions via hydrogen bonds are weaker than covalent bonds, it can be expected that atomic movements involving H-bonds have low frequency vibrational modes. Sub-Terahertz (sub-THz) vibrational spectroscopy that combines measurements with molecular dynamics (MD) computational prediction has been demonstrated as a promising approach for biological molecule characterization. Multiple resonance absorption lines have been reported. The knowledge of relaxation times of atomic oscillations is critical for successful application of THz spectroscopy for hydrogen bond characterization. The purpose of this work is to use atomic oscillations in 0.35-0.7 THz range, found from molecular dynamic (MD) simulations of *E.coli* thioredoxin (2TRX), to study relaxation dynamics of two intramolecular H-bonds, O·H–N and O·H–C. Two different complimentary techniques are used in this study, one is analysis of statistical distribution of relaxation time and dissipation factor values relevant to low frequency oscillations, and the second is analysis of autocorrelation function of low frequency quasi-periodic movements. By studying hydrogen bonds atomic displacements, it was found that the atoms are involved in a number of collective oscillations, which are characterized by different relaxation time scales ranging from 2-3 ps to more than 150 ps. The existence of long lasting relaxation processes opens the possibility to directly observe and study H-bond vibrational modes in sub-THz absorption spectra of bio-molecules if measured with appropriate spectral resolution. The results of measurements using a recently developed frequency domain spectroscopic sensor with a spectral resolution of 1 GHz confirm the MD analysis.

## 1. Introduction

Hydrogen bonds (H-bonds) in biological macromolecules are important for their structure and functions. Yet there are no simple direct methods

---

[journal], [year], [vol], 00–00 | 1

<sup>a</sup>Department of Electrical and Computer Engineering, University of Virginia, 351 McCormick Road, P.O. Box 400743, Charlottesville, VA 22904-4743, USA. E-mail: tg9a@virginia.edu

<sup>b</sup>Vibratess LLC, 104 Chaucer Rd, Charlottesville, VA 22901, USA

---

to observe and characterize H-bonds. Since interactions via hydrogen bonds are weaker than covalent bonds, it can be expected that atomic movements involving H-bonds have low frequency vibrational modes. Sub-Terahertz (sub-THz) vibrational spectroscopy of biological macromolecules, which combines measurements with molecular dynamics (MD) computational prediction, has been demonstrated as a promising approach for studying interactions between low energy radiation and intra-molecular dynamics. It reveals resonance spectroscopic features, vibrational modes or group of modes at close frequencies in absorption (transmission) spectra of biomaterials caused by a fundamental mechanism of interaction of low frequency internal intra-molecular motions via hydrogen and other weak bonds with THz radiation. Although multiple resonance absorption lines in sub-THz region have been reported in measurements with appropriate spectral resolution, for example<sup>1-6</sup>, successful application of THz spectroscopy for DNA, RNA and protein characterization requires deep understanding of relaxation processes of atomic dynamics (displacements) within a macromolecule.

The dissipation time is one of the fundamental problems related to THz vibrational modes in biological molecules. The width of individual spectral lines and the intensity of resonance features observed in sub-THz spectroscopy are sensitive to the relaxation processes of atomic movements within a macromolecule. It is clear that the decay (relaxation) time,  $\tau$ , is the factor that determines the spectral width and the intensity of vibrational transmission/absorption modes, the required spectral resolution, and eventually the discriminative capability of sub-THz spectroscopy. The suggested range of molecular dynamics relaxation times for processes without bio-molecular conformational change varies from approximately 1.5 ps to 1 ns in different studies.<sup>7,8</sup> The corresponding values for the dissipation factor,  $\gamma$ , and the width of spectral lines,  $W$ , which are reciprocal to  $\tau$ , are between 6 and 0.01  $\text{cm}^{-1}$ . Values of  $W$  above 1  $\text{cm}^{-1}$  would result in structureless sub-THz spectra, since vibrational resonances could not be resolved in the case of the high density of low intensity vibrational modes.<sup>3</sup> The existence of long-lasting dynamic processes have been confirmed by relaxation dynamics of side chains in macromolecule thioredoxin observed by time-resolved fluorescence experiments.<sup>9</sup> At the same time the entire mechanism that determines relaxation times in dynamics processes is still not completely understood. There are a number of studies on the problem using MD simulation and Langevin equation along with the analysis of inelastic neutron scattering<sup>10,11</sup> and other experimental techniques.<sup>12</sup> The estimates from inelastic neutron scattering give very large broadening of low frequency motions. Possible reasons for the differences between experiment and simulation results have been discussed in<sup>13</sup>, in particular

much higher vibrational density of states in simulations compared to neutron scattering experiments. It is known from experiments that “proteins exist in an ensemble of structures, described by an energy landscape”<sup>14</sup>, and neutron scattering spectra result from an average over different proteins conformations or substates. These motions, however, are quite different from quasi-harmonic vibrational modes in THz, and especially in sub-THz spectral range, for both, time and displacement scales.<sup>15</sup> Weak THz vibrations associate with displacements at distances on the order of only  $\sim 0.1 - 1.0$  Angstrom. These oscillations might live for a relatively long time since collision probability is less when displacements are small.

Another problem existing in analysis of relaxation is that in most cases the data on relaxation cannot be attributed to atomic fluctuations in a certain frequency range. For example, the most common method for relaxation analysis, autocorrelation function of atomic displacements, often exhibits featureless, non-exponential decays.<sup>16</sup> One explanation for this result is that atoms go through different frictional regions in space and time, and the average produces non-exponential decay.<sup>16</sup> However, another explanation may lie in the fact that the autocorrelation function usually averages oscillations over all frequencies observable in MD simulation. Regarding proteins, it has also been shown that relaxation processes can be a complex and heterogeneous phenomenon.<sup>17</sup>

In this work, statistical analysis of MD data is applied to study relaxation in vibrational dynamics of two intra-molecular H-bonds, O··H–N and O··H–C, of *E.coli* thioredoxin. These bonds connecting tryptophan with charged neighbors have been studied in<sup>9</sup> using fluorescent spectroscopy. By studying atomic displacements obtained from MD in time and frequency domains, we found that the atoms in these bonds are involved in a number of collective oscillations, which are characterized by different relaxation time scales ranging from 2-3 ps to more than 150 ps for processes without conformational change. Two different complimentary techniques are used in this study, one is analysis of statistical distribution of relaxation time (or dissipation factor) values relevant to low frequency oscillations, and the second is analysis of autocorrelation function of low frequency quasi-periodic movements.

The existence of long lasting relaxation processes makes it possible to directly observe and study H-bond vibrational modes in sub-THz absorption spectra of bio-molecules. Until recently we used a Fourier transform spectrometer Bruker IFS 66v with the spectral resolution of  $0.25 \text{ cm}^{-1}$ .<sup>3</sup> Although significant progress in experimental THz spectroscopy was demonstrated and reliable information was received for transmission/absorption spectra from different biological macromolecules and species, experimental characterization still required milligrams quantities of material, a detector cooled with liquid helium

for reliable characterization, and a system under vacuum or purged with dry gas because of the very low intensity of radiation available from the mercury lamp source and to eliminate disturbances caused by liquid and vapor water.<sup>3</sup> The implementation of THz vibrational spectroscopy was  
5 impeded because of the absence of spectroscopic systems, which simultaneously satisfy the requirements of good spectral and spatial resolution, along with high sensitivity. For our new studies, we utilize a spectroscopic sensor prototype developed by Vibratess.<sup>4</sup> This novel constant wave, frequency-domain spectroscopic instrument with imaging  
10 capabilities operates without the need for cryogenic cooling of the detector. The high sensitivity, good spectral resolution of  $0.03\text{ cm}^{-1}$ , and a spatial resolution below the diffraction limit permitted us to observe intense and narrow spectral resonances in transmission/absorption spectra of nanogram samples of biological materials with spectral line  
15 widths as narrow as  $W = \sim 0.1\text{ cm}^{-1}$ . Transmission spectra were obtained in the sub-THz region between 315 and 480 GHz for both, macromolecules and biological species. The results of measurements using this new spectroscopic sensor confirm our MD analysis.

## 2 Methods

20

### Molecular Dynamics

A detailed description of our protocol for MD simulation of protein thioredoxin using Amber 10 can be found in our recent study.<sup>18</sup> We simulated a complex of thioredoxin (PDB code 2TRX) and 8 Å shell of  
25 TIP3P water. In preparation steps, a constant volume and temperature (NVT) ensemble is used to raise the temperature to 293 K. The system is heated for  $\sim 16$  ps, and protein's atoms are restrained using a 10 kcal/mol/Å<sup>2</sup> force constant. During this heating process, bonds involving hydrogen are fixed. Constant pressure periodic boundary conditions are  
30 used to scale the system volume during 100 ps to reach a density of  $\sim 1\text{ g/cm}^3$ . In these procedures, a 10 Å real space cutoff is used with a 2 fs integration time step. Once the system has attained selected values of temperature and density, random velocities from the Maxwellian distribution are assigned to all atoms, followed by another equilibration  
35 step (NPT ensemble) for further energy minimization.

After equilibration, a 5.0 ns production run is performed in a constant volume and energy ensemble (NVE) to avoid undesirable effects on atomic motions due to coupling to an external thermal bath. During the simulation, the coordinates of all atoms of the system are  
40 recorded every 20 fs for the entire production run.

Atomic trajectories collected in MD simulations are converted to the covariance matrix of atomic displacements  $\langle R_i R_k \rangle$ . The force-field

matrix is found in a quasi-harmonic approximation utilizing the relation between the covariance matrix and the inverse of the force-constant matrix ( $\langle R_i R_k \rangle = k_B T [F^{-1}]_{ik}$ ), where  $R$  are displacements and  $F$  are force constants.<sup>19,20</sup> Diagonalization of  $F$  matrix gives eigenfrequencies (normal mode frequencies) and eigenvectors (displacement vectors-normal modes).

In our study<sup>18</sup>, MD simulations of sub-terahertz vibrational modes of the protein thioredoxin were conducted with the goals of finding the conditions needed for simulation convergence, improving the correlation between experimental and simulated absorption spectra, and ultimately for enhancing the predictive capabilities of computational modeling. We studied the consistency, accuracy and convergence of MD simulations of the sub-THz vibrational modes by comparing simulations performed using different initial conditions, protocols and parameters to the experimental results. It was demonstrated that the constant energy simulation protocol, NVE, during the production run is more preferable than the constant temperature regime, NPT, for several reasons. Using the NVE ensemble in a production run gives more stable results compared to NPT regime. Constant energy simulations without frequent exchange with the external bath for temperature regulation induce fewer disturbances into trajectories of atoms and, in addition, prevent transition of a protein molecule into a different conformation. At the same time, the NVE protocol requires more attention to the choice of starting energy in the production run. The starting total energy of the system (ETOT) at the beginning of the production run significantly affects results and has to be close to the equilibration minimum.

### Absorption Spectra Calculation

Using atomic trajectories from the constant energy and volume MD simulations, thioredoxin's sub-THz vibrational spectra and absorption coefficients were calculated in a quasi-harmonic approximation. The absorption coefficient spectra  $\alpha(\nu)$  as functions of the frequency  $\nu$  can be calculated through the relationship between  $\alpha$  and the imaginary part of dielectric permittivity<sup>21</sup>:

$$\alpha(\nu) = W \nu^2 \sum_k \frac{S_k}{(\nu^2 - \nu_k^2)^2 + W^2 \nu^2} \quad (1)$$

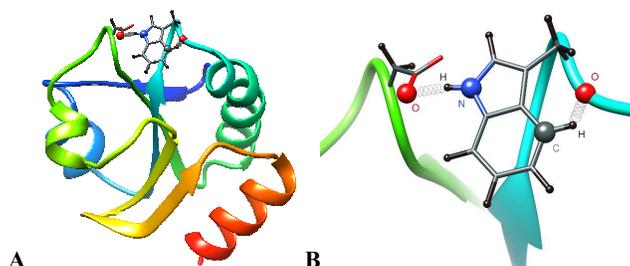
where  $\nu_k$  are normal mode frequencies calculated by diagonalization of the force-constant matrix, and  $S_k$  are oscillator strengths computed for all vibrational modes  $k$ . Two values of line width for all vibrational modes in sub-THz range were earlier suggested from our experimental work:  $W = 0.5 \text{ cm}^{-1}$  (moderate spectral resolution in Bruker spectrometer), and

$W = 0.1 \text{ cm}^{-1}$  from high resolution spectroscopy using Vibratess spectrometer.

Better simulation convergence and improved consistency between simulated vibrational frequencies and experimental data were obtained using a new procedure for averaging mass-weighted covariance matrices of atomic trajectories in MD simulations.<sup>18</sup> In particular, the open source package ptraj was edited to improve a matrix analyzing function. Averaging of only six matrices gives much more consistent results, with absorption peak intensities exceeding those from the individual spectra, and with a relatively good correlation between simulated vibrational frequencies and experimental data. Reasonably good correlation between absorption spectrum of thioredoxin simulated with  $W = 0.5 \text{ cm}^{-1}$  and experimental results as measured with a moderate spectral resolution of  $0.25 \text{ cm}^{-1}$  have been demonstrated.<sup>18</sup> Experimental spectra taken with much better spectral resolution of  $0.03 \text{ cm}^{-1}$  compared with computational modeling using  $W = 0.1 \text{ cm}^{-1}$  also demonstrated reasonable correlation.<sup>4</sup>

### H-bonds studied

Fig. 1 (A and B) shows two H-bonds in thioredoxin that we studied in this work. In one of these bonds, O··H–N, the oxygen atom (OD1, ASP61) from aspartate ASP61 (on the left of Fig.1B) has a weak interaction through the hydrogen atom (HE1, TRP31) to the nitrogen (NE1, TRP31) in the indole ring of tryptophan (TRP31). The location of this bond allows for oscillations of a relatively large (4-7 Å) amplitude.



**Fig. 1.** Two H-bonds in thioredoxin: one is an O··H–N interaction formed between Asp 61 and TRP 31 involving O (OD1, ASP61), H (HE1, TRP31) and N (NE1, TRP31) atoms. The second interaction O··H–C is formed between 4th C atom (CE3, TRP31) of the TPR31 indole ring through H (HE3, TRP31) to an oxygen atom (O, TRP31) involved in the peptide bond between TRP31 and CYS32 residues. **A:** the entire molecule is shown. **B:** detailed view of the two bonds. Images were made using Chimera.<sup>22</sup>

The second hydrogen bond C–H··O connects the 4<sup>th</sup> carbon atom (CE3, TRP31) from the same indole ring of tryptophan residue TRP31 through its hydrogen atom (HE3, TRP31) to the oxygen atom (O, TRP31) in the peptide bond between TRP31 and CYX32 (shown on the right of Fig.

1B). These two particular H-bonds have been earlier used in experiments on fluorescent spectroscopy<sup>9</sup> to study relaxation times for quenching excited tryptophan TRP31. For the first considered hydrogen bond, O··H–N, we analyzed the dynamics of distances between O and H and  
5 between O and N atoms during MD simulation. A similar approach was used for the second hydrogen bond O··H–C.

Two independent complimentary methods for data analysis were used in this study. In the first method of studying atomic displacements, we analyzed statistical distributions of spectral line width, relaxation  
10 time values, and vibrational frequencies in the sub-THz spectral range. In the second method, an analysis of autocorrelation function of low frequency quasi-periodic movements was used.

### **MD Trajectories and Fast Fourier Filtering**

15 The coordinates of chosen atoms are extracted from the MD trajectory to calculate the distance between a pair of atoms (1 and 2) for each time point in the 5 ns simulation as

$$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2} \quad (2)$$

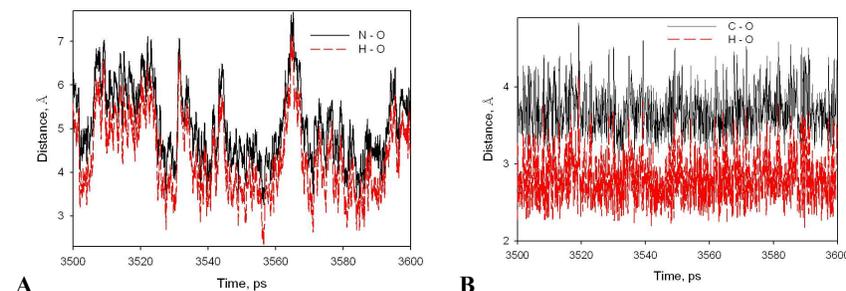
20 For better accuracy, a distance trajectory was interpolated to have a 1 fs time step using Matlab *spline* function.

Since we are interested in low energy vibrational modes in the sub-THz range, to facilitate data analysis and understanding results for fast  
25 vibrations with a very small amplitude, the distance fluctuations in time domain are processed using a Fast Fourier Transform (FFT) filter available at the Matlab website.<sup>23</sup> Application of a low frequency rejection filter with the boundary below 0.3 THz provides more uniform motion dynamics. In the FFT filtering procedure, the time domain  
30 trajectory is first transferred into the frequency domain, then frequency components inside filter limits are cut, and after that the inverse FFT procedure is applied to return the data back into the time domain.

## **3. Results**

### **Statistical distributions of relaxation parametrs using fitting to a model of two non-interacting damped oscillators**

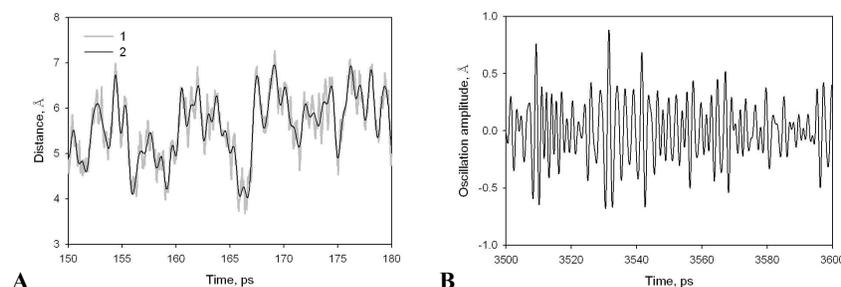
35 Figures 2 A and B show fragments of MD trajectories for two atom pairs involved in two H-bond interactions demonstrated in Fig 1.  
40



**Fig. 2.** Dynamics of the distances. **A:** O·H–N hydrogen bond interaction: solid - N and O; dash – H and O; **B:** O·H–C hydrogen bond interaction: solid - C and O; dash – H and O.

Considering the bond O·H–N in Fig. 2A, several important results become obvious immediately: 1) the atoms are involved in heterogeneous movements with different time durations and distance variations up to 3–3.5 Å, and 2) the O–N and O–H distances change with time in a similar way. This second result could be expected since the chemical bond between H and N is very strong and the distance fluctuations between them are significantly smaller compared to the distance O–H. Variations of distances between atoms in the second H-bond, O·H–C, demonstrated in Fig. 2B, are significantly smaller, only ~0.5 Å most of time, although very short variations with an amplitude of ~1 Å are also observed. The similarities between C–O and H–O distance variations are, however, preserved. A peak in the time interval 3560–3570 ps observed for H–O distance in the O·H–N bond in this particular simulation run is not revealed in the dynamics of O·H–C interaction. This peak, involving movement at a distance ~4 Å, is probably an indicator of conformational change.

Fig. 2 represents all movements involving these hydrogen bonds that might include possible vibrations at all frequencies. Since we are interested in a relatively narrow sub-THz frequency range, as a next step, we used a standard Fast Fourier Transform (FFT) filtering technique to separate interfering motions shown in Fig. 2. For example, application of a low frequency reject FFT filter removes movements with amplitudes of several Angstroms and periods of tens and hundreds of picoseconds, similar to the peak at time 3565 ps observed in Fig. 2A. Application of a reject filter above frequency of 1.2 THz removes high frequency oscillations from a raw trajectory (gray line 1, Fig 3A) and simplifies the trajectory to black line 2. Application of both, low frequency and high frequency reject filters, leaves only movements that occur in the frequency range that we are interested in, with much more uniform motion dynamics as demonstrated in Fig. 3B, with obvious quasi-harmonic movements.



**Fig. 3.** Application of FFT filter. **A:** Gray line (1): original N-H...O trajectory from MD simulation, black line (2): fluctuations with frequencies above 1.2 THz were removed using 0-1.2 THz band-pass filter. **B:** Application of a 0.35-0.7 THz band-pass filter removes both, higher frequency and lower frequency movements and reveals periodic oscillations in a sub-THz frequency range of our interest.

The existence of long-lived vibrations with different periods above 1.3 ps, which corresponds to vibration frequencies below  $\sim 26 \text{ cm}^{-1}$  is clearly demonstrated. The presence of harmonic components indicates that there are local regions in the molecule, with atoms oscillating in this frequency range for at least of several periods. These small amplitude vibrations have been described as “rattling motions in a cage consisting of the neighbouring atoms within a molecule or surrounding solvent”.<sup>24</sup>

In the case of quasi-harmonic movement, the distance between two atoms  $Y(t)$  as a function of time can be modelled as a damped oscillator:

$$Y(t) = A \cos(\omega t + \varphi) e^{-\gamma t} \quad (3)$$

where  $A$  is an amplitude ( $\text{\AA}$ ),  $\gamma$  is a dissipation factor or relaxation rate ( $\text{ps}^{-1}$ ),  $t$  is time (ps),  $\omega$  is angular frequency ( $2\pi/\text{ps}$ ), and  $\varphi$  is phase. At the same time the trajectory shown in Fig. 3B is more complicated and can not be described by a model of only one oscillator even for a relatively small time interval. The simplest possible model to describe the filtered trajectory for a quantitative analysis during a relatively short time interval is a model of two non-interacting damped oscillators:

$$Y(t) = Y_1(t) + Y_2(t) \quad (4)$$

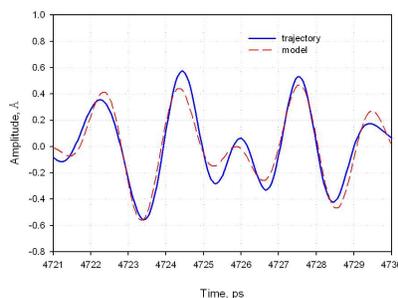
where each component is presented by eqn (3).

The fact that relatively long lived vibrations are able to survive without fast dissipation indicate that the damping factors,  $\gamma$ , for at least some vibrational modes, have to be relatively small. More detail analysis, described below, is required for identification of other motions that are definitely present and for studying accurate values of relaxation parameters.

Using a moving time window (frame)  $\Delta T$  with a  $\delta t$  step along a filtered trajectory we apply the model of two oscillators to find parameters  $A$ ,  $\omega$ ,  $\varphi$  and  $\gamma$  for each frame and each step using a fitting

procedure. Time frames  $\Delta T$  of 5, 10 or 20 ps and time steps  $\delta t$  of 50 fs or 100 fs have been used to find optimal values of parameters that give the clearest presentation of results. Matlab *lsqcurvefit* function is used to make the fitting to the filtered data. Lower and higher boundaries for relaxation rate were set to  $0.001 \text{ ps}^{-1}$  and  $10 \text{ ps}^{-1}$ , which corresponds to 0.1-1000 ps relaxation time limits. The quality of fitting is verified by calculation of the Pearson correlation ( $r$ ) between the fitting curve and the raw data, and the percent of unexplained variation is found as  $100(1-r^2)\%$ . This parameter is used as a cutting threshold to remove results with low quality fitting. Fig. 4 gives an example of fitting results in relative MD coordinates after filtering using the model inside a 10 ps window. In this particular example, the portion of unexplained variation in the data is less than 10%. Parameters of two oscillators from fitting in this example are presented in Table 1.

15



**Fig. 4.** Example of fitting the filtered trajectory of O-N distance. Solid line: MD data, dash line: model. The portion of unexplained variation in this example is less than 10%.

**Table 1.** Parameters of two oscillators for the given trajectory window in Fig. 4

oscillator	$A$ , Å	$f = \omega/(2\pi)$ , THz	$\gamma$ , $\text{ps}^{-1}$
$Y_1$	-0.26	0.588	0.003
$Y_2$	-0.34	0.372	0.045

20

To find the statistical distribution function for a dissipation factor,  $\gamma$ , all fitting results were further grouped into classes using equal intervals,  $\delta\gamma$ , and the number of observations was calculated for each class. For the purpose of comparing simulation results with experimental data, we also found the distribution function for the absorption Lorentzian line width at half peak ( $W$ ) using its relationship with the dissipation factor of a damped oscillator,  $\gamma$ :

$$W = \frac{\gamma}{\pi \cdot c} \quad (5)$$

where  $\gamma$  is in  $\text{ps}^{-1}$ ,  $c$  is the speed of light ( $\text{cm} / \text{ps}$ ), and  $W$  is in  $\text{cm}^{-1}$ . The line width and the dissipation factor differ only by a scaling factor  $\pi c$ . The procedure similar to finding  $\gamma$  was used to find the distribution function for relaxation time  $\tau = \gamma^{-1}$  using equal intervals,  $\delta\tau$ , and for frequencies of vibrations using equal intervals,  $\delta f$ . The quality of data presentation for distribution of spectral width values is sensitive to the choice of intervals, with too large of intervals resulting in a small number of possible observations, while too small interval gives noisy results. Several values of  $\delta W$  were used to find the best presentation of results. Fig. 5 demonstrates this effect for the spectral line widths distribution for the O·H–N bond.

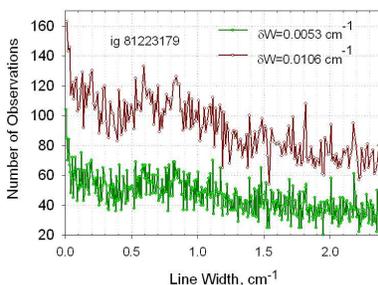
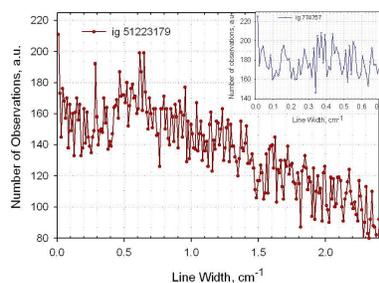
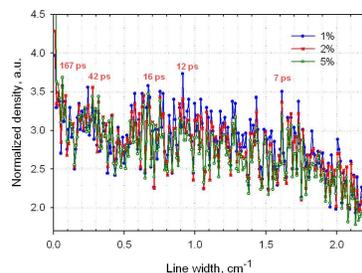


Fig. 5. Statistical distribution of spectral line widths for O·H–N bond vibrations (O–N distance) calculated with two values of  $\delta W$ ,  $0.0106 \text{ cm}^{-1}$  and  $0.0053 \text{ cm}^{-1}$ . Moving window 5 ps, filter 0.35–0.7 THz, threshold 2%.

The main result is, however, the same independent of  $\delta W$ : the distribution is not a smooth function of line width values, but shows sharp peaks. In addition to the existence of very narrow spectral lines (widths less than  $0.1 \text{ cm}^{-1}$ ) having relatively high probability, almost equal distribution of line widths up to  $W=1 \text{ cm}^{-1}$ , and slow reducing of distribution of widths above  $\sim 1 \text{ cm}^{-1}$  are clearly demonstrated. Similar results are obtained for the second hydrogen bond, O·H–C, (Fig. 6). Line width distributions shown in Figures 5 and 6 are not sensitive to the values of unexplained variations in the fitting procedure. This result is also demonstrated in Fig. 7, where the peak positions are independent of the cutting threshold.

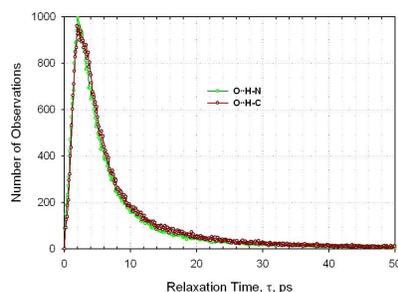


**Fig. 6.** Statistical distribution of spectral line widths for O-H-C bond vibrations (O-C distance).  $\delta W=0.0106 \text{ cm}^{-1}$ , moving window 5 ps, filter 0.35-0.7 THz, threshold 2%. The inset shows a more detailed fragment below  $0.7 \text{ cm}^{-1}$  from another MD simulation.



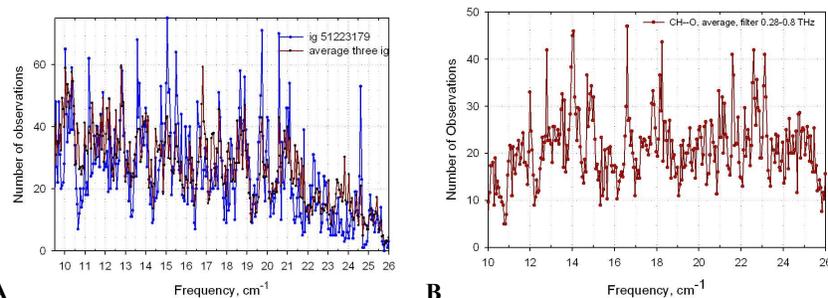
**Fig. 7.** Comparison of line width distributions calculated with different percent of unexplained variations. Parameters used: one 5 ns production run, moving time window 5 ps with 50 fs step,  $\delta W=0.0106 \text{ cm}^{-1}$ , filter 0.35-0.7 THz. Numbers above curves are relaxation time values at local peaks of line width distributions.

Relaxation times can be recalculated as  $1/\gamma$  from line width distributions in Figures 5-7 using eqn (5); some results are shown as numbers above curves in Fig. 7. Thus the atoms are involved in a number of collective oscillations, which are characterized by different relaxation time scales. However, Fig. 8 gives much more accurate results for the range of small relaxation times as calculated directly from time distribution for two hydrogen bond vibrations using filtered sub-range of 0.35-0.75 THz. The almost identical results for both bonds indicate that the distribution has a maximum at relaxation time of 2-3 ps. These small values of relaxation time have been found in different experiments, and are often attributed to rotational and translational relaxation of bulk water molecules.<sup>7,10</sup> However, relaxation time distributions show very long tail that can not be described with only one relaxation process. There is not enough sensitivity for more accurate analysis of longer relaxation time probabilities using this approach because of a small window used,  $\Delta T = 5 \text{ ps}$ .



**Fig. 8.** Distribution of relaxation time values for O·H-N and O·H-C bonds vibrations with small  $\tau$ . Filter 0.35-0.7 THz.

Fig. 9A shows the distribution of vibrational frequencies in sub-THz range. Results are sensitive to the choice of initial Maxwell velocities determined by a random seed “ig” number. Results for one ig and averaging for 3 ig numbers however definitely indicate the presence of many modes in our relatively narrow sub-range and reduced modes density above 21  $\text{cm}^{-1}$ . Vibrational frequencies are different for two bonds (Fig. 9A and 9B).



**Fig. 9. A:** Distribution of vibrational frequencies in sub-THz range for O·H-N bond using 0.28-0.8 THz filter, window 10ps, 10% threshold. **B:** The same as in A but for O·H-C bond. Simulation run with a random seed number ig 51223179.

### Autocorrelation Function

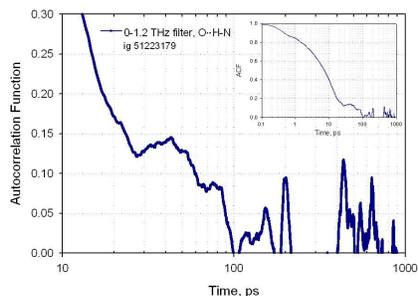
We also analyzed dissipation of THz oscillations using autocorrelation function (ACF),  $F(\Delta t)$ , that was computed as

$$F(\Delta t) = \frac{1}{N-k} \sum_{i=1}^{N-k} [(Y_{t_i} - \bar{Y})(Y_{t_{i+k}} - \bar{Y})] / \sigma_Y^2, \quad (6)$$

where  $Y_{t_i}$  and  $Y_{t_{i+k}}$  are trajectory coordinates separated by  $\Delta t = t_{i+k} - t_i = k \cdot \delta t$ . In our case, time step  $\delta t = 10$  fs, and the total number of steps,  $N =$

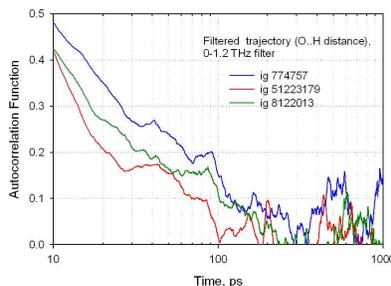
$5 \cdot 10^5$ , for 5 ns MD run.  $\bar{Y}$  and  $\sigma_Y^2$  are the sample mean and variance of  $Y$ .

Two types of time-domain data were analyzed using ACF. First, ACF was calculated for the original trajectory of distances between studied atoms. Second, it was computed for FFT filtered oscillations where upper and lower frequency limits were varied. ACF of oscillations after applying a 0-1.2 THz band-pass filter is shown in Fig. 10 for time above 10 ps, and for the entire range is given in the inset.



**Fig. 10.** ACF of oscillations after applying a 0-1.2 THz band-pass filter to the original trajectory of distances between O and N atoms of the O-H-N bond for time above 10 ps. The entire range is given in the inset.

Finally, Fig. 11 shows the variation in results for O-H-N bond ACF calculated from three MD simulations. All results reveal overall rather smooth reducing of ACF with time for  $t$  below  $\sim 30$  ps.



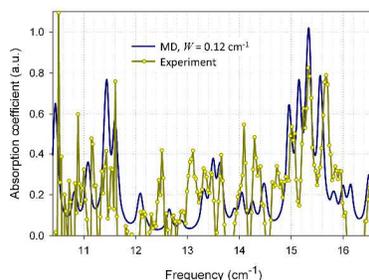
**Fig. 11.** Autocorrelation functions for the filtered trajectory of distances between O and N atoms in O-H-N bond calculated from three MD simulation runs.

However at higher time values, the peaks are reproducibly observed at  $\sim 40$ , 80-90, and  $\sim 170$  ps from all three simulations. Above  $\sim 200$  ps the results from different simulation runs are not consistent. This fact might indicate the possibility of molecule transition into a different conformational state. It can also explain the variability of other results from different simulation runs having different initial Maxwell velocities distribution.

## 4. Comparison with Experimental Data and Discussion

We have recently measured absorption spectra from *E. coli* protein thioredoxin using a new spectrometer with high spectral resolution.<sup>4</sup> The existence of intense and narrow spectral resonances was observed in transmission/absorption spectra with spectral line widths as narrow as  $W \sim 0.1 \text{ cm}^{-1}$ .

In Fig. 12, the experimental spectrum is compared with MD simulation to verify both. Although the overall correlation between the theory and experimental data confirms again the existence of intense and narrow absorption lines, not all peaks are reproduced in the measured and simulated spectra. The most obvious reason for differences is that the same damping factor  $\gamma$  was used to calculate absorption for all vibrational modes. Our current results demonstrate a broad distribution of spectral line widths (Figures 5-7) and of damping factor  $\gamma$ . Additional explanation has to be given to the fact that narrow width vibrations having small values of  $\gamma$  and large values of  $\tau$  are well observed in absorption data and, at the same time, the highest probability is demonstrated for vibrations with very short  $\tau \sim 2 \text{ ps}$  (Fig. 8). Figures 9A and 9B demonstrate vibrational frequencies for two hydrogen bonds as simulated with different spectral line widths. However, not all of these vibrations contribute equally to the absorption spectrum shown in Fig. 12. For each particular vibrational mode, the absorption coefficient is inversely proportional to  $W$  (eqn 1), and absorption due to vibrations with small relaxation times (or with large  $\gamma$  and spectral line width) has very low probability of being observed.



**Fig. 12.** Absorption spectrum of protein thioredoxin from *E. coli*: MD simulation and experimental results as measured using Vibratess spectroscopic sensor. Our work.<sup>25</sup>

## 5. Conclusions

In this work, relaxation dynamics for two intra-molecular hydrogen bonds, O·H–N and O·H–C were studied. Time domain trajectories (atomic displacements) demonstrate the complicated character of atomic

movements. Fourier transform filtering was used to restrict the data into a frequency domain within the limit of 0.35-0.7 THz, which corresponds to the experimental conditions used to characterize vibrational spectra. This procedure permitted selection and visualization of quasi-harmonic components of atomic movements. The fluctuations were further analyzed by fitting with a model of two non-interacting damped oscillators. A new fitting procedure permitted us to extract more information about relaxation parameters from atomic trajectories and justifies application of statistical approach for quantitative analysis of oscillations. In the statistical analysis used, sample distributions for relaxation factors and spectral line widths reveal multiple peaks with almost equally high probability in the range between 0.1 and 1  $\text{cm}^{-1}$ . By studying atomic displacements, it was found that the atoms are involved in a number of collective oscillations, which are characterized by different relaxation time scales ranging from 2-3 ps to more than 150 ps. The relaxation time distribution gives a sharp peak for  $\tau \sim 2$  ps, that was earlier observed in different experiments, and a very long tail that can not be described with only one relaxation process. A broad spectrum of vibrational frequencies is demonstrated. The results from analysis of autocorrelation functions for hydrogen bond distances confirm the existence of processes with long time characteristic scales.

The existence of long relaxation processes makes it possible to directly observe and study H-bond vibrational modes in sub-THz absorption spectra of bio-molecules measured with the proper spectral resolution. The results obtained in this study are in general agreement with measurements using a recently developed time domain spectroscopic sensor with the spectral resolution of 1 GHz. These new results permit better understand the dynamics of hydrogen bonds in thioredoxin and will be further used for more detailed analysis and prediction of experimental sub-THz absorption spectra from biological macromolecules.

### Acknowledgements

We would like to express our great appreciation to Dr. Ngai Wong, Senior Physical Scientist at DTRA, for his valuable support during the planning and development of this research work. This work is supported by the contract from the US Army Research Office #W911NF-12-C-0046.

### References

- 1 D. F. Plusquellic and E. J. Heilweil, Terahertz Spectroscopy of Biomolecules, in: THz Spectroscopy: Principles and Applications, Book CRC Press, 2007, 269 – 297.

- <sup>2</sup> W. Zhang, E. R. Brown, M. Rahman and M. L. Norton, *Applied Physics Letters*, 2013, **102**(2), 023701–023701–4.
- <sup>3</sup> T. Globus, T. Dorofeeva, I. Sizov, B. Gelmont, M. Lvovska, T. Khromova, O. Chertihin and Y. Koryakina, *American Journal of Biomedical Engineering*, 2012, **2**(4), 143–154.
- <sup>4</sup> T. Globus, A. M. Moyer, B. Gelmont, T. Khromova, M. I. Lvovska, I. Sizov and J. Ferrance, *IEEE Sensors Journal*, 2013, **13**(1), 72–79, 2013.
- <sup>5</sup> T. Globus, I. Sizov and B. Gelmont, *Adv. in Biosci. and Biotech.*, 2013, **4**(3A), 493–503.
- <sup>6</sup> I. Sizov, M. Rahman, B. Gelmont, M. L. Norton and T. Globus, *Chemical Physics*, 2013, **425**, 121–125.
- <sup>7</sup> T. Li, A. A. Hassanali, Y.-T. Kao, D. Zhong and S. J. Singer, *J. Am. Chem. Soc.*, 2007, **129**(11), 3376–3382.
- <sup>8</sup> K. E. Furse and S. A. Corcelli, *J. Phys. Chem. Lett.*, 2010, **1**(12), 1813–1820.
- <sup>9</sup> W. Qiu, L. Wang, W. Lu, A. Boechler, D. A. R. Sanders and D. Zhong, *Proc. Natl. Acad. Sci. U.S.A.*, 2007, **104**(13), 5366–5371.
- <sup>10</sup> K. Moritsugu and J. C. Smith, *J Phys Chem B*, 2005, **109**(24), 12182–12194.
- <sup>11</sup> J. Smith, T. Becker, S. Fischer, F. Noé, A. Tournier, G. Ullmann and V. Kurkal, *Physical and Functional Aspects of Protein Dynamics*, in: *Soft Condensed Matter Physics in Molecular and Cell Biology*, Taylor & Francis, 2006, 225–241.
- <sup>12</sup> N. Q. Vinh, S. J. Allen and K. W. Plaxco, *J. Am. Chem. Soc.*, 2011, **133**(23), 8942–8947.
- <sup>13</sup> E. Balog, J. C. Smith and D. Perahia, *Phys Chem Chem Phys*, 2006, **8**(6), 5543–5548.
- <sup>14</sup> H. Frauenfelder, B. H. McMahon, R. H. Austin, K. Chu and J. T. Groves, *Proc. Natl. Acad. Sci. U.S.A.*, 2001, **98**(5), 2370–2374.
- <sup>15</sup> S. Cusack and W. Doster, *Biophys. J.*, 1990, **58**(1), 243–251.
- <sup>16</sup> M. D. Ediger, *Annu Rev Phys Chem*, 2000, **51**, 99–128.
- <sup>17</sup> A. A. Golosov and M. Karplus, *J Phys Chem B*, 2007, **111**(6), 1482–1490.
- <sup>18</sup> N. Alijabbari, Y. Chen, I. Sizov, T. Globus, and B. Gelmont, *J Mol Model*, 2012, **18**(5), 2209–2218.
- <sup>19</sup> M. Karplus and J. N. Kushick, *Macromolecules*, 1981, **14**(2), 325–332.
- <sup>20</sup> R. M. Levy, M. Karplus, J. Kushick and D. Perahia, *Macromolecules*, 1984, **17**(7), 1370–1374.
- <sup>21</sup> T. Globus, D. Woolard, M. Bykhovskaia, B. Gelmont, L. Werbos, and A. Samuels, *International Journal of High Speed Electronics and Systems*, 2003, **13**( 04), 903–936.

- 22 E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M.  
Greenblatt, E. C. Meng and T. E. Ferrin, *J Comput Chem*, 2004,  
25(13), 1605–1612.
- 23 [http://www.mathworks.com/matlabcentral/fileexchange/12377-  
5 interactive-fourier-filter-version-1-  
5/content/FourierFilter15/FouFilter.m](http://www.mathworks.com/matlabcentral/fileexchange/12377-interactive-fourier-filter-version-1-5/content/FourierFilter15/FouFilter.m)
- 24 R. M. Daniel, R. V. Dunn, J. L. Finney and J. C. Smith, *Annual  
Review of Biophysics and Biomolecular Structure*, 2003, **32**(1), 69–  
92.
- 10 25 T. Globus, A. Moyer, B. Gelmont, I. Sizov and T. Khromova,  
Dissipation Time in Molecular Dynamics and Discriminative  
Capability of Sub-THz Spectroscopic Characterization of Biological  
Molecules and Cells, *Chemical and Biological Defense Science and  
Technology Conference*, DTRA, Las Vegas, Nov. 2011.
- 15