Invited Paper

Terahertz spectroscopy of biomolecules

Dongshan Wei [1], Zhongbo Yang [1], Mingkun Zhang [1], Mingjie Tang[1], Shihan Yan[1], Changcheng Shi[1], Liangping Xia[1], Huabin Wang [1], Tianying Chang [1, 2], Chunlei Du [1], and Hong-Liang Cui [1*, 2]

[1] Chongqing Key Laboratory of Multi-scale Manufacturing Technology, Chongqing Institute of Green and Intelligent Technology, Chinese Academy of Sciences, Chongqing, 400714, China

[2] College of Instrumentation Science and Electrical Engineering, Jilin University, Changchun, Jilin, 130061, China

[*1] Email: hcui@cigit.ac.cn

(Received August 18, 2015)

Abstract: Terahertz (THz) electromagnetic wave belonging to the frequency band from 0.1 to 10 *THz* (pundits narrow this range to 0.3-3 *THz*) has emerged as a powerful tool for investigating biomolecular systems. Since the energy level of THz wave largely coincides with that of the biomolecular low-frequency motions including vibration, rotation and translation of the molecular skeleton and that of the weak intermolecular interactions including hydrogen-bond and van der Waals etc., THz spectroscopy as a molecular detection technology has its unique advantages over some other existing ones. In the last several years, our group has focused on THz spectroscopy detection and spectral imaging of biomolecules, especially on the development of a THz near-field microscopy equipment for imaging of cells and even biomacromolecules. On the theoretical front, we have calculated and analyzed the characteristic spectra of polypeptides and investigated the effects of conformation and size of biomolecules on their THz spectra. Experimentally, microfluidic channels for THz spectroscopy tests were fabricated and the THz spectra of λ-DNA were investigated. A scattering-type scanning near-field terahertz microscopy based on the atomic force microscopy and the THz spectrometer was built and initial results about the modulation of the THz near-filed signals were presented.

Keywords: Terahertz spectroscopy, Biomolecule, Absorption coefficient, Near-field microscopy, Molecular dynamics, Microfluidic channel.

doi: 10.11906/TST.101-112.2015.09.10

1. Introduction

Terahertz (THz) wave refers to the frequency ranging from 0.1 to 10 *THz* [1]. As a biomolecular

detection technology, THz wave has some unique advantages. For instance, the energy level of THz wave largely coincides with the level of the biomolecular low-frequency motions such as vibration, rotation and translation of the molecular skeleton. These movements can be identified through fingerprint characteristics in the THz transmission or absorption spectroscopy [2, 3]. THz vibrations are fundamentally different from infrared ones which are mostly from the strong bonding role of intramolecular neighboring atoms [4]. THz radiation excites low-frequency molecular vibrations from intra/inter-molecular domains connected by the weak interactions including hydrogen bonds, van der Waals and non-bonded (hydrophobic) interactions etc. [5]. Therefore, THz spectroscopy can be applied as a unique molecular fingerprint technology. Furthermore, the energy of THz wave is relatively low with a value of ~4 *meV* per photon, which will hardly cause the ionizing damage to the structure of biomolecules [6].

Terahertz spectroscopy of biomolecules has been developing rapidly during these years. In experiment, biomolecules such as nucleic acids, nucleobases and nucleosides [7-9], short chain peptides [10, 11], amino acids [12], collagen [13], artificial RNA [14] have been characterized by THz spectroscopy technology. Nagel et al. [15] reported a resonant THz sensor to analyze double stranded DNA unwinding process. Castro-Camus et al. [16] used terahertz spectroscopy to observe the conformational changes of photoactive yellow protein. Menikh et al. [17] used a pulsed terahertz wave biosensor to monitor the binding interaction between biotin and avidin molecules. Mernea et al. [18] investigated macromolecular crowding effects on a highly concentrated bovine serum albumin solution using THz spectroscopy.

In theory, density functional theory (DFT) is widely used to calculate the THz absorption spectra of biomolecules [19-21]. Energy minimization combined with harmonic approximation based on the molecular mechanics method has also been applied to analyze low frequency vibration modes of biological molecules [3, 22, 23]. Results obtained from the above two methods have good stability and reproducibility. However, these methods can only deal with biomolecules in solid or crystalline state but not in aqueous state. Biomolecules will combine with surrounding water to form hydrogen bonds in solution. Nonpolar residues are also affected by repulsive interaction of water molecules (hydrophobic interaction). Electrostatic interaction may also exist between solute and solution molecules. All of these interactions will attribute to biomolecular motions [10]. Therefore, THz absorbance due to the hydration water around biomolecules has to be taken into account as it differs from the bulk water [24]. To simulate THz spectra of biomolecules in aqueous solution, molecular dynamics (MD) simulation is an appropriate tool. Bykhovski et al. [23] calculated the THz spectra of tRNA in implicit and explicit water by MD simulations, and found the explicit water model seems to adequately represent an effect of water in cell on tRNA dynamics.

Li et al. [22] obtained THz spectra of a DNA decamer by MD simulation and energy minimization calculation, and confirmed that MD simulation can better explain the experimental result comparing to the energy minimization calculation.

In this article, we summarized our recent theoretical and experimental work on the THz spectroscopy studies of biomolecules. Characteristic spectra of polypeptides were calculated and effects of conformation and size of polypeptides on the THz spectra were investigated by using molecular dynamics simulations. Micro/nanofluidic channels were fabricated and the THz spectra of λ-DNA were investigated by a photomixing continuous-wave THz spectroscopy combined with the channels. Initial testing results about the home-built scattering-type scanning near-field terahertz microscopy were presented.

2. Conformation and chain length dependent terahertz spectra of alanine polypeptides

Three different linear alanine polypeptides whose sequences are ACE-(ALA)n–NME with n=5, 15, 30, respectively, where ACE and NME are general termini of polypeptides, were built using ff99SB force field in AMBER 12 software [25] and were simulated at 310 *K*. Each of these polypeptides was first simulated in vacuum for 1.0 *ns*, and then was solvated in explicit water with TIP3P model and an energy minimization was carried out with a weak restraint of 5.0 $kcal/mol \bullet \mathring{A}^2$ on the center of mass of each chain. Next the restraint on the polypeptide chain was removed and an NPT (constant atom number, pressure and temperature) ensemble simulation of 2.0 *ns* was performed to adjust the system to reach a density of approximate 1.0 g/cm^3. Finally, a long NVT (constant atom number, volume, and temperature) ensemble simulation about 10 to 40 *ns* for different polypeptides with a time step of 1.0 *fs* was performed for conformation analysis.

THz spectra of polypeptides were calculated through the Kramers–Heisenberg relation, in which the absorption coefficient can be denoted as

$$\alpha(\omega) \sim \omega^2 \sum_k \frac{S_k \gamma_k}{\left(\omega^2 - \omega_k^2\right)^2 + \gamma_k^2 \omega^2} ,$$

where ω_k is vibration eigenfrequency, S_k is the oscillator strength and γ_k is oscillator dissipation factor. ω_k can be determined from diagnosing the covariance matrix of the coordinate fluctuations according to the quasi-harmonic approximation and S_k is proportional to the square of the dipole moment of the polypeptide molecule. Because the frequency band of interest involved in our discussion is narrow, the decay can be considered to be frequency independent approximately

and $\gamma_k \equiv 1.0\ cm^{-1}$ in all calculations.

THz spectra of alanine polypeptides in coiled conformation at 310 *K* are shown in Fig. 1a. It can be seen that absorption becomes stronger with the increase of wave number. The resonance features indicate that THz spectra are significantly dependent on chain length. The average absorption intensity of three chains was plotted in Fig. 1b. It can be seen that the average absorption intensity increases with the increase of the chain length. To explain the reason of stronger absorption in the longer polypeptide chain, the number of hydrogen bonds formed by polypeptides with different chain lengths and water molecules were counted and shown in Fig. 1c. Here the hydrogen bond is defined with a maximum donor-acceptor distance of 3.5 *Å* and a minimum donor-H-acceptor angle of 135 °. From Fig. 1c, we can see the average number of hydrogen bonds increases from 13.9 to 55.6 for polypeptides from Ala_5 to Ala_{30}. More hydrogen bonds will strengthen the turbulence of water on the polypeptide chain, thus the absorption intensity is enhanced.

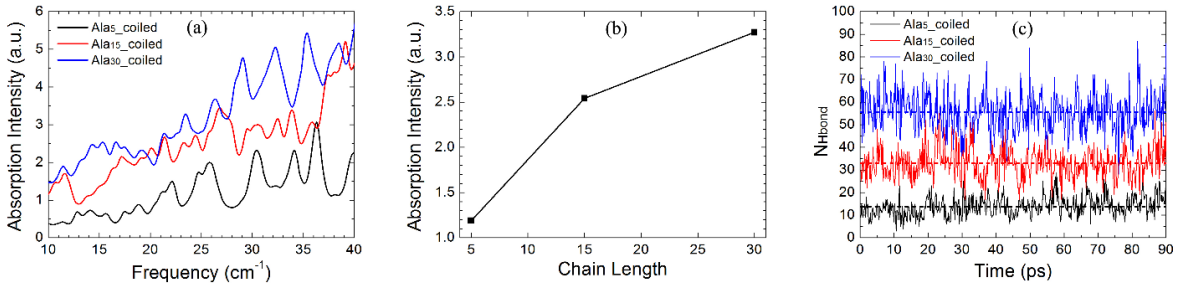

Fig. 1(a) THz spectra of polypeptides with different chain lengths. (b) The averaged absorption intensity of three polypeptides. (c) Variations of the number of hydrogen bonds for polypeptides, the dash line represents the average number of hydrogen bonds for each polypeptide.

Coiled and helical spectra of Ala_{15} are compared in Fig. 2a. The average absorption intensity in coiled conformation is higher than that in helical one. This result is consistent with the experimental finding in Ref. [26], where the unfolded lysozyme (coiled) has a stronger absorption compared with the native state (helical). To explain the reason of stronger absorption in coiled conformation, similarly, the number of hydrogen bonds formed by coiled or helical Ala_{15} and water molecules was calculated and shown in Fig. 2b. The average number of hydrogen bonds is 33.0 in coiled conformation and 26.4 in helical one. THz spectra of biomolecules in solution are mainly caused by vibrations and torsions from intra-molecular domains, and weak bonded resonance absorptions from inter-molecular domains. However, the absorption coefficients of biomolecules themselves in aqueous are relatively small with respect to water, so the resonance absorptions of biomolecules themselves are mostly buried [27], then THz spectra mainly reflect the inter-molecular interactions of weak bonds, mainly including hydrogen bonds and van der Waals bonds, but the oscillator

strength of van der Waals bond is an order of magnitude smaller than the hydrogen bond, so the absorption peaks of THz spectra mainly relate to the dynamics of the interacting Hydrogen bonds. H-bonds vibrating with different frequencies will form different absorption peaks and vibrating with the same frequency will enhance the intensity of the absorption peak.

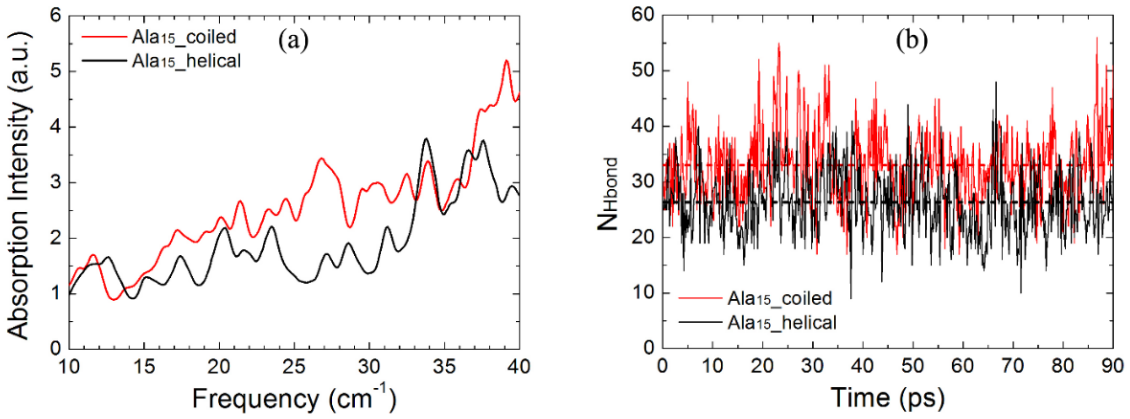

Fig. 2(a) THz spectra of Ala_{15} in different conformations. (b) Variations of the number of hydrogen bonds in two conformations, the dash line represents the average number of hydrogen bonds for each conformation.

Except for the effect of the interacting H-bonds on the THz absorption spectra of Ala_{15} with different conformations, the flexibility of the overall molecule structure may also affect the absorption peaks and absorption intensity as reported by Heyden et al. [27]. They investigated the effects of protein flexibility on the THz spectra of protein and found that THz absorption intensity decreases when the ubiquitin folds and becomes rigid, which is consistent with our conclusion that absorption of helical Ala_{15} (rigid) is smaller than that of coiled Ala_{15} (flexible). Therefore, the differences of the THz absorption spectra between the coiled and helical polyalanines are determined by two factors: the interacting H-bonds and the structural flexibility, but the former affects more than the latter since the intrinsic absorption of biomolecules themselves is much weaker than that of water.

3. Terahertz absorption signal of dissolved DNA detected in a microfluidic chip

A micro-channel array chip with quartz plates as the substrate was fabricated and used to detect terahertz (THz) absorption signals of λ-DNA solution. The microscopic image of one of the micro-channel array structures is shown in Fig. 3. The periodic line width of the fabricated chips may vary from 2 to 50 μm, depending on the size of the detected biological targets. The depth of the

micro-channels is fixed to be 2 μm. Fabrication processes of the micro-channel array chip including: 1) Produce the micro structure pattern on mask plate via laser direct writing lithography; 2) Spin coating photoresist on quartz plate, then transfer the pattern on mask to quartz substrate via interferometric lithography; 3) Etch the structure to target depth using reactive ion etching, then wash the residual photoresist to fully clean the substrate.

Fig. 3 Optical micrograph from top side of a 25 μm pattern microfluidic chip.

THz spectroscopy of λ-DNA was tested by a photomixing continuous-wave THz spectroscopy combined with the chip. The spectrum is obtained by sweeping the single-frequency THz transmit photomixer over a specified frequency range, while coherently converting the signal transmitted through the sample under test with a receiver photomixer. The frequency resolution is 0.35 *GHz* and the dynamic range is 79 *dB* at 500 *GHz*. The sweep time from 0.6 to 1.2 *THz* is about 13 minutes. λ-DNA (Thermo company) was dispersed in TE buffer in a mass concentration of 0.3 $\mu g/\mu L$.

To perform the THz spectrum measurement, 10 μL TE buffer was first injected into the channel with a pipette, whose THz spectrum is recorded as the reference signal, $P_B(v)$. Then the chip was washed clearly and 10 μL DNA solution was injected and its spectrum is recorded as the sample signal, $P_S(v)$. Last the THz path was blocked and the noise floor was measured as $P_N(v)$. Variations of the reference, sample and noise signals with the frequency are shown in Fig. 4. As can be seen, three obvious absorption peaks of water vapor locate around 750 *GHz*, 985 *GHz*, 1160 *GHz*, which is consistent with the results of Roggenbuck et al. [28]

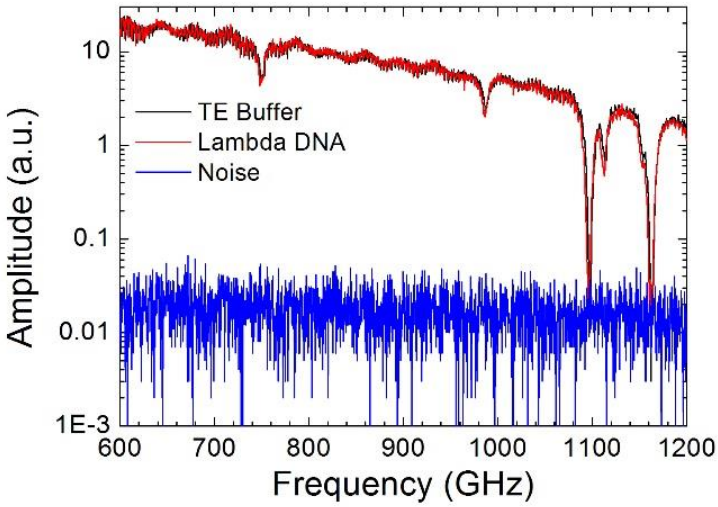

Fig. 4 The spectra of TE buffer, λ-DNA solution and noise floor.

Three independent measurements of DNA solution were carried out and the relative transmission of dissolved DNA was calculated by the ratio of the DNA signal to the buffer signal, i.e., $\mathrm{T}(v) = \mathrm{P}_S(v)/\mathrm{P}_B(v)$. Fig. 5 showed the average relative transmission of λ-DNA at the frequency from 600 ~ 1100 *GHz*. It can be seen clearly that there are three absorption peaks appearing at 635, 850 and 930 *GHz* and all of these peaks have a narrow FWHM less than 20 *GHz*. The result is in good agreement with Brown's experiment result [29], where obvious absorption peaks at 850 and 950 *GHz* were found for λ-DNA in EDTA buffer.

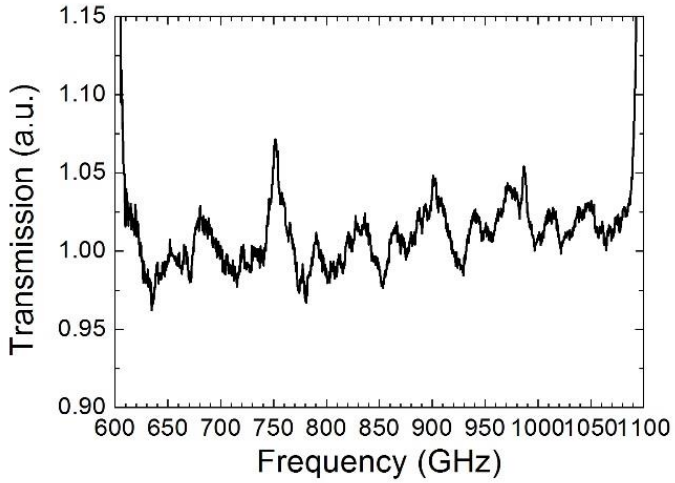

Fig. 5 The averaged relative transmission of λ-DNA.

4. Tests of scattering-type scanning near-field terahertz microscopy

THz spectroscopy has great applications in detecting molecules and studying the carrier dynamic of semiconductor. However, the spatial resolution of conventional THz spectroscopy is limited by diffraction to about a half wavelength and thus it is incompetent to detect a single biomacromolecule (typically a few nanometers) and image a semiconductor nanodevice. To obtain nanoscale spatial resolution, a home-built scattering-type scanning near-field terahertz microscopy (S-SNTS) based on an atomic force microscopy (AFM) and a THz spectrometer is built and under testing in our laboratory (Fig. 6). In this setup, a metalized AFM probe as a broadband antenna confines free-space propagating terahertz waves into a nano-scale enhanced electric field near the tip apex of probe, and simultaneously scatters the near-field terahertz signal of sample to the far-field. The spatial resolution of the S-SNTS is approximately equal to the curve radius of the tip (typically 20 ~ 30 *nm*), so only a few, even a single molecule under the tip apex can be detected and imaged. The scattering terahertz intensity of the tip apex is inversely proportional to the fourth power of the curve radius of the tip due to the fact that the size of tip apex is much shorter than the wavelength of the incident terahertz wave.

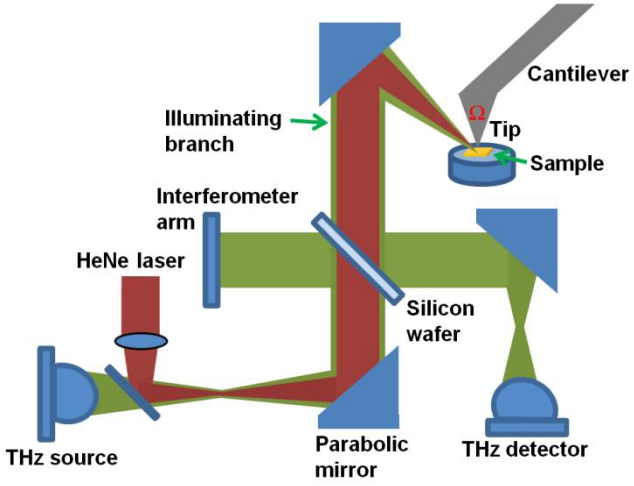

Fig. 6 The experimental setup of S-SNTS.

A 3-*mm* thickness high-resistance silicon wafer is used as terahertz beam splitters to split the collimated terahertz waves into two beams with a reflectivity to transmissivity ratio of around 54:46. One beam of terahertz waves is vertically incident on the interferometer arm equipped with an Al-coated reflector and another beam is illuminating on the metalized probe. First, the probe was replaced by the Al-coated reflector, completely reflecting the terahertz wave in the illuminating branch. As showed in Fig. 7a, a similar feature with periodic peak valleys on the transmission spectroscopy of both the interferometer arm (black curve) and the illuminating branch (red curve),

respectively was observed by blocking either branch. The frequency difference of two adjacent peak valley ($\Delta\nu$ = 14.8 *GHz*) is consistent with the theoretical calculation result of 14.3 *GHz* due to the interference of terahertz waves from the multiple reflections at the silicon-air interface considering the 3-*mm* thickness silicon wafer with a refractive index of 3.418 and an incidence angle of 45 °. An additional peak valley with the frequency difference of 8.1 *GHz* in interference spectroscopy (blue curve) is the destructive interference between the interferometer arm and the illuminating branch. As shown in Fig. 7a, the amplitude of the interference spectroscopy is larger than those of the interferometer arm and the illuminating branch at the same frequency, indicating that the Michelson-type interferometer works well. In Fig. 7b, it can be seen that the signal scattering from the large shaft and cantilever (black curve) is about three times of that from the "U" stainless steel AFM head and metal substrate surface (red curve). The subsequent work is to verify the obtained near-field terahertz signal of the sample by vibrating the cantilever at its mechanical resonance frequency and employing the lock-in amplifier to detect the terahertz signal. Detection of cells and biomacromolecules like DNA and protein using the scattering-type scanning near-field terahertz microscopy is around the corner.

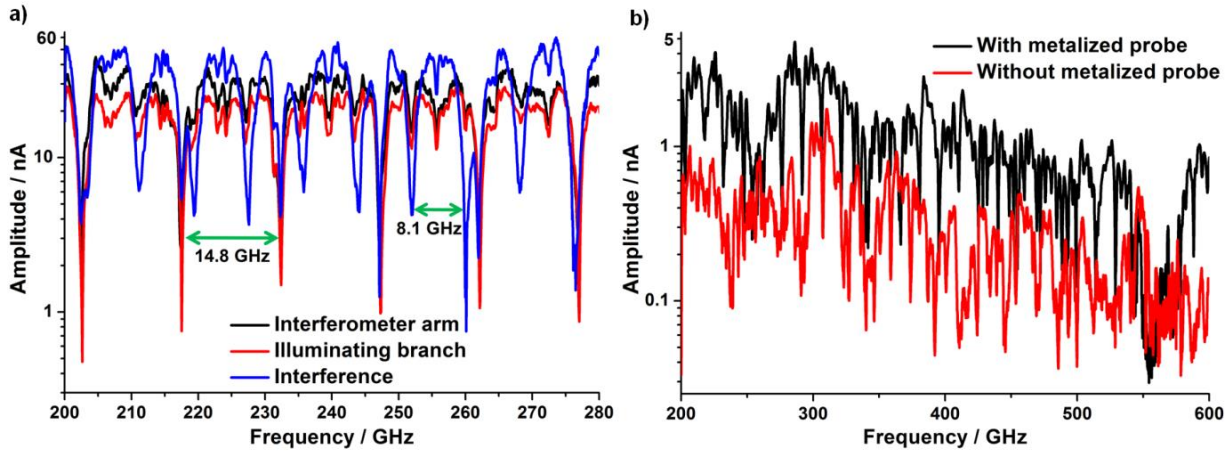

Fig. 7(a) Transmission signals of interferometer arm, illuminating branch and the interference between them. (b) The spectral response of the metalized probe.

5. Conclusions

We have reported the recent progresses on the THz spectroscopy detection of biomolecules from theoretical and experimental aspects. In theory, owing to differences in the size, conformation and the intrinsic flexibility, biomolecules have different THz absorption/transmission spectra, which

can be well discerned by molecular dynamics simulations. In experiment, due to the strong absorption of water, THz spectra of biomolecules in solution is difficult to be accurately detected. Adopting the high-resolution photomixing continuous-wave THz spectroscopy and the micro-channel array chip, characteristic absorption peaks of λ-DNA are measured. The scattering-type scanning near-field terahertz microscopy is built with the object for imaging of cells and biomacromolecules. Present testing results suggests our work is on the way to develop a near-field terahertz microscopy.

ACKNOWLEDGMENTS

This work is part supported by National 973 Program of China (No. 2015CB755401), fundamental & advanced research project of Chongqing, China (cstc2013jcyjC00001), national natural science foundation of China (Nos. 21407145 and 31400625), application development project of Chongqing, China (cstc2013yykfC00007), and scientific equipment research project of Chinese Academy of Sciences (Development of THz imaging spectrometer for biomacromolecules).

References

[1] K. Sakai, and M. Tani. “Introduction to terahertz pulses”. *Terahertz Optoelectronics,* 97, 1-30(2005).

[2] T. Chen, Z. Li, and W. Mo. “Identification of biomolecules by terahertz spectroscopy and fuzzy pattern recognition”. *Spectroc. Acta Pt. A-Molec. Biomolec. Spectr*., 106, 48-53(2013).

[3] H.L. Zhang, E. Zukowski, R. Balu, et al. “A dynamics study of the A-chain of ricin by terahertz vibrational calculation and normal modes analysis”. *J. Mol. Graph*., 27, 655-663(2009).

[4] B.H. Stuart, and D.J. Ando. “Biological applications of infrared spectroscopy”. *John Wiley & Sons*, (1997).

[5] T. Globus, T. Dorofeeva, I. Sizov, et al. “Sub-THz vibrational spectroscopy of bacterial cells and molecular components”. *American Journal of Biomedical Engineering*, 2, 143-154 (2012)

[6] S.W. Smye, J.M. Chamberlain, A.J. Fitzgerald, et al. “The interaction between terahertz radiation and biological tissue”. *Phys. Med. Biol*., 46, 101-112(2001).

[7] B.M. Fischer, M. Walther, and P.U. Jepsen. "Far-infrared vibrational modes of DNA components studied by terahertz time-domain spectroscopy". *Phys. Med. Biol*., 47, 3807-3814(2002).

[8] Y.C. Shen, P.C. Upadhya, E.H. Linfield, et al. "Temperature-dependent low-frequency vibrational spectra of purine and adenine". *Appl. Phys. Lett.*, 82, 2350-2352(2003).

[9] Y.C. Shen, P.C. Upadhya, E.H. Linfield, et al. "Observation of far-infrared emission from excited cytosine molecules". *Appl. Phys. Lett.*, 87, 011105(2005).

[10] T. Ding, A.P.J. Middelberg, T. Huber, et al. "Far-infrared spectroscopy analysis of linear and cyclic peptides, and lysozyme". *Vib. Spectrosc*., 61, 144-150(2012).

[11] T. Ding, T. Huber, A.P.J. Middelberg, et al. "Characterization of low-frequency modes in aqueous peptides using far-infrared spectroscopy and molecular dynamics simulation". *J. Phys. Chem. A*, 115, 11559-11565(2011).

[12] T.M. Korter, R. Balu, M.B. Campbell, et al. "Terahertz spectroscopy of solid serine and cysteine". *Chem. Phys. Lett.*, 418, 65-70(2006).

[13] A.G. Markelz, A. Roitberg, and E.J. Heilweil. "Pulsed terahertz spectroscopy of DNA, bovine serum albumin and collagen between 0.1 and 2.0 THz". *Chem. Phys. Lett*., 320, 42-48(2000).

[14] B.M. Fischer, M. Hoffmann, H. Helm, et al. "Terahertz time-domain spectroscopy and imaging of artificial RNA". *Opt. Express*, 13, 5205-5215(2005).

[15] M. Nagel, F. Richter, P. Haring-Bolivar, et al. "A functionalized THz sensor for marker-free DNA analysis". *Phys. Med. Biol*., 48, 3625-3636 (2003).

[16] E. Castro-Camus, and M.B. Johnston. "Conformational changes of photoactive yellow protein monitored by terahertz spectroscopy". *Chem. Phys. Lett*., 455, 289-292(2008).

[17] A. Menikh, S.P. Mickan, H.B. Liu, et al. "Label-free amplified bioaffinity detection using terahertz wave technology". *Biosens. Bioelectron.*, 20, 658-662(2004).

[18] M. Mernea, O. Calborean, L. Petrescu, et al. "Protein association investigated by THz spectroscopy and molecular modeling". Laser Applications in Life Sciences 2010, International Society for Optics and Photonics, , 73760O-73769 (2010).

[19] G. Wang, and W.N. Wang. "Experimental and theoretical investigations on the terahertz vibrational spectroscopy of alanine crystal". *Acta Physico-Chimica Sinica*, 28, 1579-1585 (2012).

[20] Z.P. Zheng, and W.H. Fan. “First principles investigation of L-alanine in terahertz region”. *J. Biol. Phys*., 38, 405-413(2012).

[21] F. Zhang, O. Kambara, K. Tominaga, et al. “M. Hayashi, Analysis of vibrational spectra of solid-state adenine and adenosine in the terahertz region”. *Rsc Advances*, 4, 269-278 (2014).

[22] X.W. Li, T. Globus, B. Gelmont, et al. “Terahertz absorption of DNA decamer duplex”. *J. Phys. Chem. A*, 112, 12090-12096(2008).

[23] A. Bykhovski, T. Globus, T. Khromova, et al. “An analysis of the THz frequency signatures in the cellular components of biological agents – art”. no. 62120H, in: D.L. Woolard, R.J. Hwu, M.J. Rosker, J.O. Jensen (Eds.) Terahertz for Military and Security Applications IV, Spie-Int Soc Optical Engineering, Bellingham, H2120-H2120(2006).

[24] S. Ebbinghaus, S.J. Kim, M. Heyden, et al. “An extended dynamical hydration shell around proteins, Proceedings of the National Academy of Sciences of the United States of America”. 104, 20749-20752(2007).

[25] D.A. Case, T.E. Cheatham, T. Darden, et al. “The Amber biomolecular simulation programs”. *J. Comput. Chem*., 26, 1668-1688(2005).

[26] T. Globus, T. Khromova, R. Lobo, et al. “THz characterization of lysozyme at different conformations”. in: Defense and Security, International Society for Optics and Photonics, 54-65(2005).

[27] M. Heyden, and M. Havenith. “Combining THz spectroscopy and MD simulations to study protein-hydration coupling”. *Methods*, 52, 74-83(2010).

[28] A. Roggenbuck, H. Schmitz, A. Deninger, et al. “Coherent broadband continuous-wave terahertz spectroscopy on solid-state samples”. *New J. Phys*., 12, 13(2010).

[29] E. Brown, E. Mendoza, Y. Kuznetsova, et al. “THz signatures of DNA in nanochannels under electrophoretic control”. in: SENSORS, 2013 IEEE, IEEE, 1-3(2013).