Spectroscopy 24 (2010) 153–158
DOI 10.3233/SPE-2010-0423
IOS Press

Low-frequency dynamics of biological molecules studied by terahertz time-domain spectroscopy

Shintaro Kawaguchi [a], Ohki Kambara [b], Carlito S. Ponseca Jr. [b], Mikihiro Shibata [c], Hideki Kandori [c] and Keisuke Tominaga [a,b,*]

[a] *Graduate School of Science, Kobe University, Nada, Kobe, Japan*
[b] *Molecular Photoscience Research Center, Kobe University, Nada, Kobe, Japan*
[c] *Nagoya Institute of Technology, Showa-ku, Nagoya, Japan*

Abstract. By terahertz (THz) time-domain spectroscopy we have measured low-frequency spectra of amino acid (glycine; Gly), short peptides ($(Gly)_3$ and $(Gly)_4$), six globular proteins and bacteriorhodopsin (BR). From the analysis of the THz spectra we defined and obtained the reduced absorption cross sections for these cases, which are proportional to the vibrational density of states. We observed anharmonic behaviors in the low-frequency modes of the short peptides. The globular proteins we investigated show a universal feature in the low-frequency spectra. BR shows the dynamical transition in the temperature dependence of the THz spectrum when the sample is hydrated.

Keywords: Terahertz time-domain spectroscopy, globular protein, bacteriorhodopsin

1. Introduction

Terahertz frequency region (THz; 1 THz $\approx$ 33 cm^{-1}) is important for biology. First, small molecules relevant to biology such as amino acids or nucleobases have a hydroxyl group or amino group, which forms intermolecular hydrogen bonds. Formation and dissociation of the hydrogen bond plays a key role for structural change of the biologically important macromolecules such as proteins and DNA. The characteristic frequencies of the intermolecular hydrogen bond mostly exist in the THz frequency region. Therefore, characterization of the spectra observed for small molecules such as amino acids in the THz frequency region is a fundamental issue for biology.

Secondly, when proteins express their functions, large conformational changes often occur. These conformational changes result from collective motions of a large number of atoms. Such motions of proteins have characteristic frequencies in the low-frequency region below a few tens of wavenumbers. Furthermore, it is well known that when proteins express their functions water molecules trapped internally and those surrounding the proteins play an important role. Owing to the dramatic progress in techniques of generation and detection of freely propagating THz radiation, spectroscopic investigation in the FIR has been widely conducted on condensed phase systems including biological samples [6].

* Corresponding author: Dr. Keisuke Tominaga, Graduate School of Science, Kobe University, Nada, Kobe, 657-8501, Japan. E-mail: tominaga@kobe-u.ac.jp.

In this work, by THz time-domain spectroscopy (TDS) we have measured the low-frequency spectra of amino acids and their short peptides to see the dependence of the THz on the chain length. Furthermore, the low-frequency spectra of globular proteins and bacteriorhodopsin (BR) were measured. We investigated effect of hydration and temperature on the low-frequency spectra of BR and discuss the relation with expression of their function.

2. Theoretical background

An absorption cross section $\sigma(\tilde{\nu})$ is related to a total dipole moment of the sample, $\mathbf{M}(t)$, through the following equations [4],

$$\sigma(\tilde{\nu}) = \frac{4\pi^3 N_A}{3\varepsilon_0 h n(\tilde{\nu})}\tilde{\nu}(1 - \mathrm{e}^{-\beta hc\tilde{\nu}})I(\tilde{\nu}), \tag{1}$$

$$I(\tilde{\nu}) = \frac{1}{2\pi N}\int_{-\infty}^{\infty} \mathrm{d}t\, \mathrm{e}^{-\mathrm{i}2\pi c\tilde{\nu}t}\langle \mathbf{M}(0)\cdot\mathbf{M}(t)\rangle, \tag{2}$$

$$\mathbf{M}(t) = \sum_{i=1}^{N} \boldsymbol{\mu}_i(t). \tag{3}$$

$I(\tilde{\nu})$ is the lineshape function defined as Fourier transform of the time-correlation function (TCF) of the total dipole moment, divided by N (the number of molecules in the sample). The lineshape function is represented as the normalized quantity per one molecule. $\boldsymbol{\mu}_i(t)$ is the individual dipole moment of one molecule, which includes induced dipole moment. $n(\tilde{\nu})$ is the refractive index of the sample. The other symbols have usual meanings. The absorption cross section is, therefore, proportional to the product of the lineshape function and the frequency factor dependent on temperature, $\tilde{\nu}(1 - \mathrm{e}^{-\beta hc\tilde{\nu}})$. The TCF of $\mathbf{M}(t)$ carries information on dynamics such as low-frequency intermolecular and intramolecular vibrations, orientational relaxation of the dipoles, fluctuation of induced-dipole moments and so on. The weak interactions play crucial roles in chemical reactions, protein functions and biological activities.

To eliminate the thermal factor $(1 - \mathrm{e}^{-\beta hc\tilde{\nu}})$ in the absorption cross section $\sigma(\tilde{\nu})$ expressed by Eq. (1), we define the reduced absorption cross section (RACS) as [7]

$$\begin{aligned}\sigma_R(\tilde{\nu}) &\equiv \frac{\beta hc\tilde{\nu}}{(1 - \mathrm{e}^{-\beta hc\tilde{\nu}})} n(\tilde{\nu})\sigma(\tilde{\nu}) \\ &= \frac{2\pi^2 c N_A}{3\varepsilon_0 N}\beta\tilde{\nu}^2 \int_{-\infty}^{\infty} \mathrm{d}t\, \mathrm{e}^{-\mathrm{i}2\pi c\tilde{\nu}t}\langle \mathbf{M}(0)\cdot\mathbf{M}(t)\rangle.\end{aligned} \tag{4}$$

The RACS is considered as a product of the infrared (IR) activity and vibrational density of states (VDOS). If we assume that the IR activity is constant in the low-frequency region because such polymers do not have any specific symmetry, the RACS is proportional to the VDOS. According to the Debye theory of three-dimensional crystals, in which the motions of the constituent atoms are considered to be a collection of harmonic oscillators, the VDOS is proportional to the square of the frequency if the wavelengths of the normal modes in the crystal are long compared to the atomic spacing. If the exponent of the power-law of the RACS is smaller than two, it suggests that the low-frequency modes do not behave harmonically, instead, they are coupled anharmonically with each other.

3. Experimental procedures

The details of the THz-TDS apparatus was already reported [2]. A mode-locked Ti:sapphire laser centered at 800 nm with a pulse duration of 120 fs and a repetition rate of 80 MHz pumped a pair of photoconductive switches used as a THz emitter and a detector. A computer controlled delay was utilized for detecting the temporal waveform of the THz wave. The whole system is placed in a box with flowing dry air. By the THz-TDS system we can measure the THz spectra in the frequency range from about 5–70 cm^{-1}. The measurements were done at room temperature (20°C) if it is not specified. For the temperature dependence experiment, the sample was dipped in the liquid nitrogen to make it cold. Immediately after taking the sample out from the liquid nitrogen, the sample was placed in the THz-TDS system to start the measurement. The temperature of the sample was measured by a thermometer attached to the metal holder for the sample.

4. Results and discussion

4.1. Amino acids and short peptides

Figure 1(a) shows the absorption spectra of amino acid glycine (Gly), and peptides $(Gly)_3$ and $(Gly)_4$. The effective frequency regions, in which the signal-to-noise ratio is relatively good and the spectra can be discussed quantitatively, are below 60 cm^{-1}, 47 cm^{-1} and 50 cm^{-1} for Gly, $(Gly)_3$ and $(Gly)_4$, respectively. The absorption plot of Gly showed a band at 50 cm^{-1}, while that for $(Gly)_3$ is observed at 40 cm^{-1}, and that of $(Gly)_4$ is at 38 cm^{-1}. It can be seen that the absorption band seem to shift towards the lower frequency as the chain is increased. These bands observed for the peptides below 40 cm^{-1} are probably caused by intramolecular vibrations in $(Gly)_3$ and $(Gly)_4$, suggesting a flexible nature around the bond between the two amino acids in the chain. Since polypeptides are much longer polymers than our samples here, it can be deduced that the polymerization causes additional intensities in the absorption spectrum that may have brought about by rather weak bond between the amino acids.

Figure 1(b), on the other hand, shows log plots of the calculated RACS spectra of the three samples, normalized in terms of one mole of the amino acid or amino residue. The frequency region for the fitting is from 5 cm^{-1} to 23 cm^{-1}, where the intermolecular modes are supposed to dominate the spectral component. Glycine has an exponent of 2.16 ± 0.06, which is almost equal to two, suggesting that its low-frequency vibrational modes behave harmonically. However, for $(Gly)_3$ and $(Gly)_4$, the exponent is decreased to be 1.69 ± 0.04 and 1.43 ± 0.04, respectively, signifying that the modes are coupled anharmoincally with the other molecules.

4.2. Globular proteins

We investigated the THz spectra of several globular proteins. The proteins were used without further purification or treatment. The spectra of the RACS were measured for proteins predominantly containing α-helix structures (albumin and myoglobin), proteins predominantly containing β-sheet structures (trypsin and trypsin inhibitor), and proteins containing both α-helix and β-sheet structures (lysozyme and lactoferrin), respectively. For all the proteins investigated, no distinct sharp band was observed in the frequency range from 3 cm^{-1} to 60 cm^{-1}. All the absorption spectra increase monotonically with frequency. It is interesting to note that all the proteins have the similar intensity of the RACS per amino residue. All the spectra are approximately proportional to the power of the wavenumber in the region

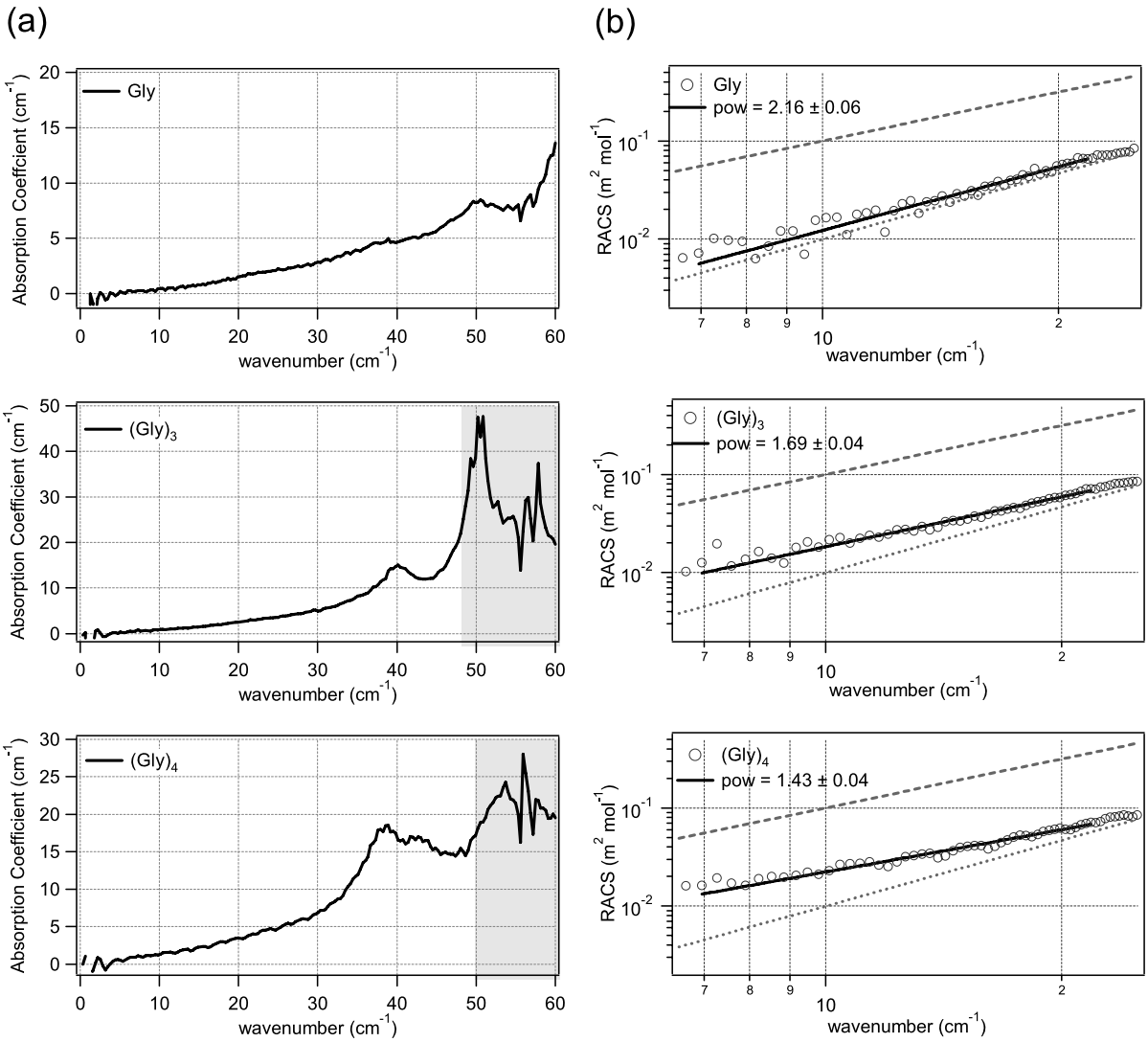

Fig. 1. Absorption spectra (a) and log plots of the RACS (b) of glycine (Gly; upper) and its peptide $(Gly)_3$ (middle) and $(Gly)_4$ (bottom). Dotted and dashed traces are the RACS of monomer and polymer of Gly reported in [7]. The frequency region where the signal-to-noise ratio is worse is shown by grey shaded region.

from 7 cm^{-1} to 55 cm^{-1}. In Table 1 we summarize the values of the exponent of the RACS for all the proteins we investigated. The values are between 1.5–1.8.

Such a universal feature of the low-frequency spectrum of globular proteins was discussed in several theoretical studies [1,5]. ben-Avraham found that the VDOS follows a characteristic, universal curve, which reflects main structural similarities among proteins [1]. The VDOS was obtained by a classical normal modes analysis based on a quadratic approximation of the potential energy of the molecule. The size of the proteins investigated ranges from 39 amino acid residues to 375 residues. The data seem to collapse into a universal curve. This scaling reflects the statistical similarities among different

Table 1
Exponent values for a power-law behavior in the reduced absorption cross sections for several proteins

Protein	Exponent[a]
Albumin	1.63 ± 0.01
Myoglobin	1.70 ± 0.02
Trypsin	1.58 ± 0.01
Trypsin inhibitor	1.60 ± 0.02
Lysozyme	1.67 ± 0.01
Lactoferrin	1.72 ± 0.02

[a]The fitting region is from 7 cm^{-1} to 55 cm^{-1}.

proteins [1]. The results of this work clearly show an experimental evidence of this universal feature of proteins.

4.3. Bacteriorhodopsin

We have measured THz spectra of the absorption coefficient and refractive index of bacteriorhodopsin (BR) by THz-TDS. One of the key issues in this experiment is to prepare a uniform sample. We prepared the sample by drying the BR solution and pressing it to make a disk of sample. From the obtained THz spectra we calculated RACS as shown in Fig. 2(a). It was found that the RACS of the BR samples shows a power-law behavior ($\sigma_R(\tilde{\nu}) \propto \tilde{\nu}^{\alpha}$). At room temperature, the exponent of the power-law of the dry sample α is estimated to 1.97 ± 0.02 from the spectral fitting in the frequency region from 7 cm^{-1} to 26 cm^{-1}. We next investigated the hydration effect on the exponent. To hydrate the sample, the dried BR sample was placed in a box in which humidity is controlled by placing appropriate saturated salt solutions. The amount of hydration is obtained by measuring the sample weight. The value of the exponent α becomes smaller as the amount of hydration increases; $\alpha = 1.74 \pm 0.02$ and 1.51 ± 0.01 for $h = 0.123$ and 0.316, respectively, where h is the weight in grams of the hydrated water per one gram of BR. For harmonic oscillators of a three-dimensional crystal lattice, its VDOS is proportional to a square of frequency. By comparing the ideal case, anharmonic coupling among the low-frequency modes of BR becomes larger as the amount of hydration increases.

Furthermore, the temperature dependence of the exponent is similar for both the dry and hydrated samples in the temperature range from -100°C to -40°C as shown in Fig. 2(b). However, above -40°C the hydrated samples show stronger temperature dependence than the dry samples. It shows that for the hydrated sample anharmonic coupling is induced above -40°C by increasing temperature. This change is due to the dynamical transition that was reported by the study of inelastic neutron scattering experiment [3]. The mechanism of the dynamical transition is following. There are multiple quasi-stable structures of protein in its free-energy surface. Below the transition temperature localized motions trapped in the quasi-stable structures are dominant, whereas above the transition temperature the interconversion between the quasi-stable structures are thermally activated, which consequently leads to the large-amplitude motions in the protein. Some proteins manifest their functions only above the dynamical transition temperature, suggesting that the dynamical transition is strongly related to the protein functions. Furthermore, this transition is only observed for the hydrated samples, and not observed in the dried sample. Therefore, hydration waters around the protein play important role in the dynamical transition.

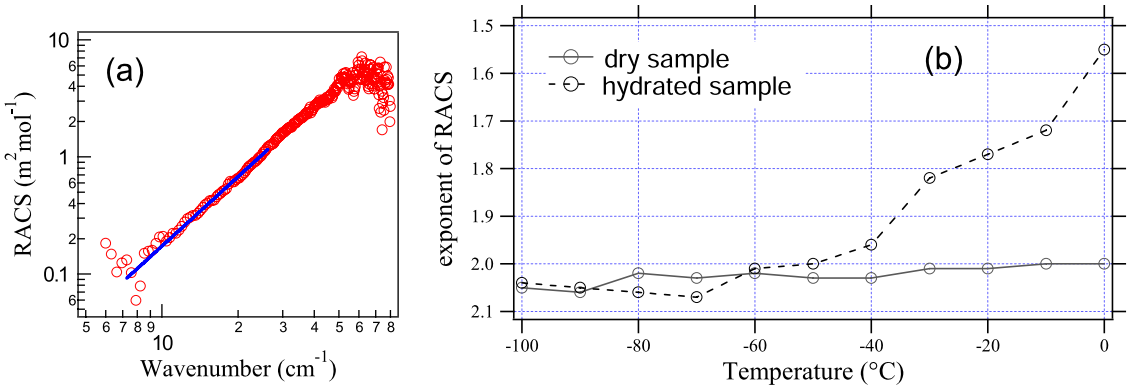

Fig. 2. (a) Reduced absorption cross section of dried bacteriorhodopsin at room temperature. The line is the result of the power-law fitting. (b) Temperature dependence of the exponent of the RACS for dried and hydrated BR.

5. Summary

In this work we have measured THz spectra of amino acid (glycine; Gly), short peptides ($(Gly)_3$ and $(Gly)_4$), six globular proteins such as albumin and bacteriorhodopsin (BR) by time-domain spectroscopy. By increasing the chain length of the peptide, we observed anharmonic behaviors in the THz spectra. The 6 globular proteins we investigated show a universal feature in the low-frequency spectra, which was predicted theoretically before. We studied the temperature dependence of the THz spectrum of BR, and we found that the dynamical transition occurs at around 230 K when it is hydrated, which is consistent with the results of inelastic neutron scattering experiment.

Acknowledgement

This work was supported by a Grant-in-Aid for Creative Scientific Research (No. 17GS0209) from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

References

[1] D. ben-Avraham, Vibrational normal-mode spectrum of globular proteins, *Phys. Rev. B* **47** (1993), 559–560.
[2] P. Dutta and K. Tominaga, Obtaining low frequency spectra of acetone dissolved in cyclohexane by terahertz time-domain spectroscopy, *J. Phys. Chem. A* **113** (2009), 8235–8242.
[3] M. Ferrand, A.J. Dianoux, W. Petry and G. Zaccai, Thermal motions and function of bacteriorhodopsin in purple membranes: Effects of temperature and hydration studied by neutron scattering, *Proc. Natl. Acad. Sci. USA* **90** (1993), 9668–9672.
[4] D.A. McQuarrie, *Statistical Mechanics*, University Science Books, Sausalito, 1976.
[5] M. Nollmann and P. Etchegoin, Universal low-frequency vibrations of proteins from a simple interaction potential, *Phys. Rev. E* **60** (1999), 4593–4596.
[6] K. Sakai (ed.), *Terahertz Optoelectronics*, Springer-Verlag, Berlin, 2005.
[7] K. Yamamoto, K. Tominaga, H. Sasakawa, A. Tamura, H. Murakami, H. Ohtake and N. Sarukura, Terahertz time-domain spectroscopy of amino acids and polypeptides, *Biophys. J.* **89** (2005), L22–L24.