ABSTRACT

THz spectroscopy is a relatively new spectroscopic method in which the THz useful range of frequencies is from 300 cm⁻¹ down to 3 cm⁻¹, a spectral region unachievable by FTIR spectroscopy. Multiple resonances in the absorption spectra of organic molecules exist in the THz range around 0.1–10 THz (3 to 300 cm⁻¹) where these resonances arise from electromagnetic wave interaction of the THz source with the low-frequency, and hence large-scale, vibrational modes within the organic macromolecules. These major vibrational modes correspond to the quaternary or overall structure of macromolecules. The concerted motion of protein structure can also be investigated by THz spectroscopy, which enables both energy and momentum transfer at the picosecond time scale. The short bursts of far-IR radiation in an ultra-fast time scale in THz spectroscopy allow researchers to study the dynamic processes of various materials for a large variety of applications. This review paper discusses the probability of utilizing THz spectroscopy as a probe to determine the intermolecular structures of macromolecules as well as basic technical knowledge in THz spectroscopy.

(KEYWORDS: Terahertz, spectroscopy, vibrational, proteins, organic molecules)

INTRODUCTION

Infrared (IR) spectroscopy is a classical chemical tool dating back to the early 20th century (Hagen & Rubens, 1902), whereas THz spectroscopy have only garnered serious attention about 20 years ago (Beard et al, 2002; Fattinger & Grischkowsky, 1989). The effort to study THz spectroscopy was triggered by the challenge to produce and detect ultra-short electric transients (changes in electrical voltage) as they were propagated along an electric transmission line (Beard et al, 2002). These electrical pulses create very small bursts of electromagnetic radiation in the far-IR frequency wavelength from 0.2 to 2.0 THz (6.66 to 66.6 cm⁻¹) (Fattinger & Grischkowsky, 1989). THz spectroscopy, sometimes referred as T-ray spectroscopy (Markelz, 2008), allows direct interrogation of complex permittivity of samples spanning a broad bandwidth of 0.1–8.0 THz (3.33–266.4 cm⁻¹) (Shen et al, 2003) without the need for cryogens and simultaneously allowing extremely flexible sample control and optics arrangements (Markelz, 2008).

THz rays provides an safer alternative to x-rays due to its non-ionizing radiative property for imaging purposes through cloth, skin, paper and many others imaging materials. The investigation of the THz fields has considerably grown with much emphasis concentrating on the THz spectroscopy of biological materials. Several investigations that have linked THz spectroscopy to the biological or biochemistry fields have been conducted by various groups of researchers worldwide, including (i) the analysis of low-frequency molecular vibration of DNA and RNA strands (Fischer et al, 2005; Mickan et al, 2002), (ii) several types of proteins including lysozyme, myoglobin, bacteriorhodopsin and virus-like particles (Falconer et al, 2010; Markelz et al, 2002; Png et al, 2009; Zakaria et al, 2011; Zhang & Durbin, 2006), (iii) the interaction between electromagnetic radiation and biological materials (Doria et al, 2004; Pickwell et al, 2004; Smye et al, 2001), (iv) the detection of potentially harmful biological agents (Brown et al, 2004) and, (v) the application of THz imaging to detect skin cancer (Pickwell et al, 2004; Woodward et al, 2003).

THE TECHNICAL ASPECT OF THz SPECTROSCOPY

The rapid growth of interest in the THz field is largely due to the emergence of the technology needed to access the THz region, including the development of THz source and detectors in recent years (Plusquellec et al, 2007). Vanexter and co-workers developed the THz photoconductive transmitter and detector based on silicon technology (Vanexter & Grischkowsky, 1990) spanning the frequency range of 0.1–2 THz (3.33–66.6 cm⁻¹). In the last decade, photoconductive antennae technology was manipulated and engineered...
for the generation and detection of THz rays (Fattinger & Grischkowsky, 1989; Grischkowsky et al, 1990). Most THz sources relied on inorganic semiconductor materials—gallium arsenide (GaAs), zinc telluride (ZnTe), and indium arsenide (InAs). These semiconductors can be used to transform an ultra-short (around 100 fs) near infrared pulsed laser to a THz pulse by excitation of electron-hole pairs in a semiconductor structure (Davies et al, 2002). The higher photon energy of the near-infrared pulse over the semiconductor band gap (energy difference from the top of valence band to the bottom of the conduction band) will create electron hole-pairs close to the surface of the THz source. A THz pulse is generated by the photo-excited acceleration of electron-hole pairs of the semiconductor material when the near-IR laser pulse reaches the semiconductor surface (Davies et al, 2002).

The channel of field effect transistors may be utilized as resonators for plasma waves where the dimensions and gate length of of the plasma frequency of the resonator may reach the THz range (Knap & Dyakonov, 2013). For THz pulse detection, GaAs (Lu et al, 1998), silicon (Si), and gallium nitride (GaN) (Kachorovskii & Shur, 2008) are typically used as plasma (THz) wave detectors. The THz detection mechanism is as follows (Dyakonov & Shur, 1996): incoming THz radiation excites the plasma wave in the FET that is translated as a voltage drop across the FET structure. This voltage drop between the drain to the source of the FET is correlated with a resonance response from the incoming electromagnetic radiation. The electromagnetic radiation will then provide the spectra needed to study the intra-molecular characteristic being studied where the theoretical calculations will be discussed in the following paragraphs.

Figure 1. Schematics of a FET utilized as THz detector (upper image), and its corresponding equivalent circuit (lower image) (Knap & Dyakonov, 2013)

In general, THz data measurements are obtained in the time domain (raw data) by recording the temporal field strength of a short electromagnetic pulse that lasts only for a few picoseconds. The Fast Fourier Transform (FFT) is the common algorithm used to convert these time domain signals (i.e., an estimate of the impulse response of the sample) to the frequency domain (Fischer, 2005b; Yin et al, 2007). From this conversion, the frequency response of the sample can be estimated. A THz pulse will be modified after propagating through a sample (dielectric material), because of the absorption and dispersion in the sample. By comparing this THz pulse with an appropriate reference pulse, an estimate can be obtained of the average absorption (absorption of EM waves by the sample) and the refraction (dispersion of light in the sample) from the sample measurements (Fischer 2005).
Absorptions of electromagnetic radiation in the THz region are frequently related to hydrogen-bonding interactions between separate molecules and variations in the crystal lattice (Fischer et al., 2007). THz data were acquired in the time domain by recording the temporal field strength of the short electromagnetic pulse. The following signal processing analysis is based on the THz studies of Bernd M Fischer (Fischer, 2005b).

The time delay signal, \( j(t) \) where \( t \) represents the time delay of a convolution of the electric pulse of the THz pulse \( E(t) \) with the THz detector pulse \( D(t) \) as given:

\[
j(t) = \frac{1}{T} \int_{0}^{T} E(t)D(t - \tau) \, dt \quad [1]\]

The linear response of a material to a THz ray is determined by the dielectric properties of the material where the complete information of the dielectric function of that material is contained in its reflected THz pulse (Fischer, 2005b). Consider that an electromagnetic plane wave of frequency \( \omega \) propagating through a sample in the \( z \)-direction with the dielectric constant of \( \tilde{n} = n + ik \) will produce a time dependent electric field as follows:

\[
E(z, t) = E_0(t) \exp \left\{ i\left(\omega t + \frac{\tilde{n}\omega}{c}z\right) \right\} = E_0(t) \exp \left\{ i\left(\frac{n\omega}{c}z\right) \right\} \exp \left\{ -\frac{k\omega}{c}z\right\} \quad [2]
\]

Here, the waveform before propagation through the sample is represented by the time dependent term \( E_0(t) = E_0 e^{i\omega t} \).

A THz reference pulse propagating through vacuum is given by:

\[
E_{ref} = E_0(\omega) \exp \left\{ i\omega \frac{d}{c} \right\} \quad [3]
\]

Consider now a THz pulse after propagating through a dielectric material of thickness \( d \). A THz pulse will be modified after propagating through a sample (dielectric material), because of the absorption and dispersion in the sample. Thus, the frequency dependent field after the THz pulse propagation through a material is given by:

\[
E_{sam}(\omega) = E_0(\omega) \exp T \left\{ \frac{n(\omega)}{c} \right\} d, \exp \left\{ -\frac{\kappa(\omega)}{c} d \right\} \quad [4]
\]

Where \( T \) is the Fresnel losses at the interface of the sample (Bea & Teich, 1974):

\[
T = \frac{4n(\omega)}{(n(\omega)+1)^2} \quad [5]
\]

By dividing this THz sample pulse with an appropriate THz reference pulse, an expression for the frequency dependent ratio (complex refractive index) can be obtained:

\[
\frac{E_{sam}(\omega)}{E_{ref}(\omega)} = A e^{i\Phi} = \frac{4n(\omega)}{(n(\omega)+1)^2} \exp \left\{ -\frac{\alpha(\omega)}{2} + i\frac{\omega}{2\pi} \right\} \quad [6]
\]

The absorption of light by an optical medium can be quantified by a parameter called the absorption coefficient, \( \alpha \); the index of refraction, \( n \), can be simplified as the ratio of the velocity of light in free space to the velocity of light in a medium. The estimation of the average absorption coefficient (energy propagation through homogenous system) can be derived from the imaginary part of Equation [3.6], as shown in the equation below:

Absorption coefficient, \( \alpha(\omega) = -\frac{1}{d} \ln\left(\frac{n(\omega)+1}{2n}\right)^2 A \quad [7] \)

where \( A \) is the complex amplitude.

The index of refraction (change in direction of wave propagation) can be derived from the real part of Equation [3.6] as shown in Equation [3.8]:

\[
n(\omega) = 1 + \frac{\Phi c}{\omega d} \quad [8]
\]

A schematic diagram of the THz spectroscopy system is illustrated in Figure 2. Measurements are typically conducted in a nitrogen-purged environment to reduce water vapor absorption by the THz radiation (Fischer, 2005a). An ultra-short (typically around 100 fs) laser pulse from the visible or near-IR laser (around 800 nm) is divided into 2 directions by a beam splitter (Picometrix, 2005). One laser pulse is directed to a THz emitter, which transforms the incident laser pulse into a THz pulse; the other pulse serves as a detector monitoring the temporal shape of the radiated THz waveform (Walther et al, 2002). A sample is deposited in the propagation path of the THz wave between the THz emitter and the detector. The sample is irradiated by the THz wave from the THz emitter, and the transmitted THz pulse is detected by the THz detector. The detected signal from the THz detector is then transformed by Fourier transform and the complex valued spectra of the THz wave is obtained (Fischer, 2005a). Far-IR optical properties (absorption coefficient and index of refraction) of the sample can be determined as a function of frequency from the resolved THz spectra (Ferguson & Zhang, 2002).
THz spectroscopy has several notable advantages over other forms of spectroscopic methods. Vibrational modes in the THz region are more widespread and include motions of all of the atoms in the molecular structure, whereas vibrational modes in the mid-IR wavelength are typically located within specific sites in the molecular conformation (Plusquellic et al, 2007). THz radiation has a longer wavelength than the near infra-red radiation, thus it will provide sharper images with better spatial resolution because it is less affected by Rayleigh scattering than near-infrared radiation (Beard et al, 2002). One of the many advantages of THz spectroscopy over far-IR spectroscopy is that far-IR spectroscopy uses blackbody radiation sources and phonon detectors, while THz spectroscopy incorporates bright sources and sensitive detectors that can provide a higher signal-to-noise ratio and a wider dynamic range than far-IR spectroscopy (Grischkowsky et al, 1990). Another advantage of THz spectrometry is that it allows time-resolved far-IR spectroscopy to cover the spectral range of 0.1–20 THz (3–600 cm\(^{-1}\)) in an ultra-fast time scale that provides sub-picosecond temporal resolution (Beard et al, 2002). Vivid changes in nucleic acid and component protein properties were observed when they were inserted into chromosomes or functional membranes (Sauer, 1995), which might be analysed using the picosecond time scale technology of THz spectroscopy.

There has been a notable increase in the interest in exploring the low-frequency collective vibrational modes within large biomolecules because these modes might provide valuable information about the conformational state of the biomolecules (Brandt et al, 2008). These low-frequency collective structural vibrational modes consisting of concerted large-scale motions of biomolecules dictated by biomolecule architecture were believed to bear crucial roles towards biomolecular function and conformational modification changes (Markelz, 2008). However,
spectroscopic studies in the low-frequency spectral region were often limited by technological constraints, such as insufficient light source power and detector sensitivity (Fischer, 2005b), owing to which it was difficult to resolve both energy and momentum transfers in the picosecond time scale (He et al, 2011b). THz spectroscopy with its inherent ability to probe the intra-molecular structure of materials in the low frequency spectral region. 1–8.0 THz (3.33–266.4 cm$^{-1}$) might fill in this gaping research gap.

**ANALYSIS OF ORGANIC MOLECULES BY THZ SPECTROSCOPY**

Advancements in THz technology have enabled studies by THz spectroscopy for the characterisation of very far-IR vibrational modes of chemical compounds (Ueno & Ajito, 2008). This intermolecular vibrational mode includes low-frequency vibrations around the hydrogen bonds (Walther et al, 2002), phonon modes that are IR-active intermolecular modes in a crystalline sample (Siegrist et al, 2006), and low-frequency delocalised modes that consist of the samples’ entire vibrational motions (Giraud et al, 2003). THz spectroscopy provides the means to actively investigate the quaternary structure of organic molecules by the interaction of THz radiation with low-frequency vibrational modes of organic molecules. The dielectric functions of the organic molecules benzoic acid and its monosubstituted derivatives 2-, 3- and 4-hydroxybenzoic acid in pressed pellet form were investigated at 0.5–4.5 THz (17–150 cm$^{-1}$) using THz time domain spectroscopy by Walther and co-workers at the University of Freiburg, Germany. Despite similarities in the molecular structure, the absorption spectra of these biomolecules had very little in common (Walther et al, 2002).

The high sensitivity of THz vibrational modes to small changes in the overall biomolecular structure in addition to the hydrogen bonding environment were evidenced again in a THz study of the enantiomeric crystalline structure of L-, D-, and DL-alanine (Yamaguchi et al, 2005) and the pure enantiomers and racemic compounds of tartaric acid (Fischer et al, 2007). Recently, the absorption spectra of organic crystals diglycine lithium nitrate, triglycine zinc chloride, and diglycine hydrobromide have shown clear resonant behavior in the permittivity diagram with sharp absorption peaks at 32.11 cm$^{-1}$, 44.16 cm$^{-1}$, and 51.72 cm$^{-1}$ (Trzebiatowska- Guszowska et al, 2014).

Comparisons of hydrated and dehydrated tri-alanine samples caused calculative differences in its respective THz spectra, demonstrating that use of a hydrogen-bonded solvent has a huge effect on the intermolecular structure of organic molecules investigated in the THz spectral range (Siegrist et al, 2006). A study of the intermolecular characteristics of chlorobenzene in the THz region conducted by Fischer (2007) showed that at a higher temperature (295 K), the absorption spectrum of chlorobenzene was swamped by large water absorption. In contrast, at a lower temperature (100 K), the constitution of chlorobenzene changed from liquid to crystalline; hence, a sharp absorption peak could be observed in the THz spectrum of crystalline chlorobenzene samples (Fischer et al, 2007). The effects of water molecules in THz measurements were evidenced again by (Zhang & Durbin, 2006), who reported significant differences in the THz absorption spectra of a myoglobin solution at various water concentrations, and (Liu & Zhang, 2006), with investigation of D-glucose monohydrate with THz-TDS. Liu and co-workers showed that the absorption peak of D-glucose monohydrate decreases at a higher sample dehydration rate (Liu & Zhang, 2006) as shown in Figure 3.
Figure 3. D-glucose monohydrate showcase escalating THz absorption peaks upon longer heating time, indicating that THz spectroscopy may be used as an analytical tool to study the dehydration of polycrystalline hydrate (Liu & Zhang, 2006)

THz SPECTROSCOPY AND INTERMOLECULAR PROTEIN STRUCTURE

The wide diversity in protein structure is matched by the vast variation in protein function in which each protein was specifically designed to cater to its role (Jackson & Mantsch, 1995). The context in which biochemical processes occur in proteins is crucial to the thorough understanding of their structures and functions. The relationship between a protein’s unique intermolecular structure and its corresponding function should be studied to gain a thorough view of proteins as a whole. THz spectroscopy is a relatively new method employed for studying intermolecular protein structure compared to other techniques, such as FTIR spectroscopy (Chirgadze & Nevskaya, 1976), small-angle light scattering (Aymard et al, 1999), and Raman spectrometry (Clark et al, 1981). Investigations of biomolecular dynamics and structure can be greatly aided by a vibrational spectra study of the biomolecules at the THz frequency wavelength. The concerted motion of protein structure can be investigated by THz spectroscopy, which enables both energy and momentum transfer at the picosecond time scale. Comparison between the resolved THz spectra of the heme protein and its calculated vibrational modes showed that a relationship exists between the calculated structural collective modes and the hydration dependence of the absorption coefficient spectra (He et al, 2011a). Walther and co-workers investigated 3 different retinal isomers that exist in the photoactiveproteins rhodopsin, bacteriorhodopsin, and isorhodopsin of at 0.3–3 THz (10–100 cm⁻¹) at 10 K and room temperature (Walther et al, 2000). The crystalline forms (prepared in polyethylene pellets) of these proteins differ only in their conformational structure. Clear differences between the 3 different retinal isomers’ absorption spectra, specifically 1.2–2.1 THz (40–70 cm⁻¹) at 10 K, illustrate the capability of this spectral region to distinguish isomeric samples.

The ability of THz spectroscopy to probe into the intermolecular structure of biomolecules has attracted the attention of many research groups worldwide. Markelz and co-workers at the University of Connecticut investigated the ability of THz time domain spectroscopy (THz-TDS) to identify biomolecular species, conformational states, and mutations in bacteriorhodopsin (Markelz et al, 2002). Bacteriorhodopsin samples were subjected to various temperatures and environmental conditions to alter their architectural structures. The THz-TDS spectra of bacteriorhodopsin demonstrated remarkable differences before and after structural alternation (Markelz et al, 2002) at a frequency of 1–1.5 THz (33.3–50 cm⁻¹). The phenomena observed by
Markelz and co-workers were probably due to conformational changes in the biomolecular structure correlating with a notable shift in the frequency spectrum (Markelz et al., 2002). This finding implies that the THz-TDS provides a non-destructive method for quantifying the intermolecular changes in biomolecules. It is also very apparent that THz-TDS has a high level of sensitivity to detect miniscule transformations in the conformation and mutation of biological systems.

The sensitivity of THz measurements to different \( \beta \)-pleated sheet orientations of crystalline tri-alanine and the effect of a solvate environment have been investigated by Siegrist (2006) at a frequency of 0.06–3 THz (2 cm\(^{-1}\) to 100 cm\(^{-1}\)). The X-ray structure obtained from the crystallography analysis of the anti-parallel tri-alanine showed a different structure of \( \beta \)-pleated sheet formation compared to that of parallel tri-alanine; this finding is in agreement with the THz spectra at 4.2 K where the spectra of parallel and anti-parallel tri-alanine showed distinctive spectral features (Siegrist et al., 2006). The finding by Siegrist and co-workers demonstrated the sensitivity of the THz spectral region for detecting conformational differences in secondary structures, the \( \beta \)-pleated sheet orientation in particular. The THz infrared absorption spectra lies in the range of 33.3–732.6 cm\(^{-1}\) (1–22 THz), in which a dense set of absorption features can be observed from the measurements of solid-state amino acids in this region (Kutteruf et al., 2003).

The investigation of the hydration effect on a hen egg white lysozyme (HEWL) sample by THz spectroscopy has been conducted by Markelz and co-workers in the frequency wavelengths of 0.15–1.95 THz (5–65 cm\(^{-1}\)). A very similar study on the HEWL hydration rate was also conducted by Moeller and co-workers. In their study, they used far-IR spectroscopy with a synchrotron light source in the frequency range of 0.5–1.35 THz (15–45 cm\(^{-1}\)) (Moeller et al., 1992). Both studies have confirmed that at a higher temperature, an increase in the absorption intensities was observed due to high water absorption (Knab et al., 2006; Moeller et al., 1992). At a higher hydration rate, the absorption spectra of the lysozyme sample closely resembled that of pure water (Moeller et al., 1992). The confirmation of low-frequency sample vibrational motions was validated by calculations of a dielectric relaxation model (Knab et al., 2006). Recent study utilizing the THz wavelength that compares the spectra of HEWL at its natural and fibrillar state showed that the spectra of amorphous (HEWL in natural state) and long, elongated fibrils (HEWL in fibrillar state) have significant broad spectral features that might be due to the quaternary structure of the protein itself (Zakaria et al., 2011) as shown in Figure 4.

![Figure 4](image-url)

**Figure 4:** Absorption spectra of lysozyme fibril (— — —) and monomeric lysozyme (——). Note that the spectra diverged after 150 cm\(^{-1}\) and onwards which indicate that the THz spectral region is able to differentiate the intra-molecular structures of proteins (Zakaria et al., 2011)
CONCLUSIONS

Low frequency stretching and bending vibrations, crystalline phonon vibrations, hydrogen bond stretches and torsional vibrations of crystalline structures can be investigated in the THz wave-length (Beard et al, 2002). These attractive properties of THz spectroscopy couples with the fact that the non-ionizing, able to transmit through almost any materials of THz rays enable intra-molecular structure analysis in the low frequency spectral region, a spectral domain unachievable by other spectroscopic methods. There is an increasing interest in analysing low-frequency collective vibrational modes in biomolecules since these modes may provide information about their conformational state (Fischer et al, 2007). However there are several limitations in studying biomolecular structure in the THz range. The limitations are a stable and high energy light source connected to the spectrometer is necessary to enable analysis of protein structure in the very low frequency region which is unachievable in a conventional bench-top spectrometer. This bright light source may provide a higher signal-to-noise ratio and a wider dynamic range than traditional far-IR spectroscopy (Grischkowsky et al, 1990). The second factor to be considered was that the sample has to be rigorously characterized prior to spectroscopic measurements and it is best to measure the sample in pressed pellet form to ensure consistency in the resulting spectra. Protein samples immersed in water or solutions will induce interference to the spectra in the form of emergence of rotational water lines and poor signal-to-noise ratio.

ACKNOWLEDGEMENTS

I would like to thank Professor Anton Middelberg, Dr. Robert Falconer and Associate Professor Andrew Bradley for fruitful discussions towards the completion of this manuscript.

REFERENCES


Letters 54: 490-492


