Terahertz dielectric study of bio-molecules using time-domain spectrometry and molecular dynamics simulations

Sushko, Oleksandr

Terahertz dielectric study of bio-molecules using time-domain spectrometry and molecular dynamics simulations

by

Oleksandr Sushko

Submitted to the University of London in partial fulfilment of the requirements of the Degree of Doctor of Philosophy

School of Electronic Engineering and Computer Science

Queen Mary London

United Kingdom

June 2014

Statement of Originality

I, Oleksandr Sushko, confirm that the research included within this thesis is my own work or that where it has been carried out in collaboration with, or supported by others, that this is duly acknowledged below and my contribution indicated. Previously published material is also acknowledged below.

I attest that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge break any UK law, infringe any third party's copyright or other Intellectual Property Right, or contain any confidential material.

I accept that the College has the right to use plagiarism detection software to check the electronic version of the thesis.

I confirm that this thesis has not been previously submitted for the award of a degree by this or any other university.

Signature: Oleksandr Sushko

Date: 19/03/2014

Details of collaboration and publications:

Most of the material in this thesis is based on the following publications and conference contributions:

JOURNAL PAPERS

1) O. Sushko, K. Shala, R. Dubrovka, R. Donnan, "A revised metrology for enhanced accuracy in complex optical constant determination by THz-TDS", *J. Opt. Soc. A*, Vol. 30, Issue 5, pp. 979-986, 2013;

2) O. Sushko, R. Dubrovka, R. Donnan "Terahertz spectral domain computational analysis of hydration shell of proteins with increasingly complex tertiary structure" *J. Phys. Chem. B,* 2013, 117, 16486−16492.
3) O. Sushko, K. Shala, R. Dubrovka, R. Donnan "Analysis of the numerical thickness determination of materials studied by THz-TDS" *J. Phys. Conf. Series*, 472 (2013) 012005.
4) O. Sushko, R. Dubrovka, R. Donnan "The Siren Effect: Sub-THz spectroscopy reveals that proteins influence the properties of water at greater distances than previously detected" submitted to *J. Phys. Chem.*
5) A. Atrashchenko, O. Sushko, N. Ulin, R. Dubrovka, R. Donnan, and V. Evtikhiev "Electrical properties of nanoporous matrix GaAs in terahertz frequency domain", in preparation, to be submitted to *Physica Status Solidi*

CONFERENCE PAPERS

1) R. F. Dubrovka, O. Y. Sushko, T. M. Loftus, B. Yang, K. Shala, R. S. Donnan, "Recent Advances in THz Science at Queen Mary University of London" Proc. *of ICATT'13 Conference,* Odessa, Ukraine.
2) O. Sushko, R. Dubrovka, R. S. Donnan, "Interpretation of THz absorption spectra of dissolved proteins using molecular dynamics simulations", *IRMMW-THz Conference*, September 2013, Mainz, Germany;
3) O. Sushko, R. Dubrovka, R. S. Donnan, "Interpretive modeling of THz response of bio-molecules as filter circuits", *IRMMW-THz Conference*, September 2012, Wollongong, Australia;
4) O. Sushko, R. Dubrovka, R. S. Donnan "Possibilities and challenges of material parameter extraction by THz time domain spectroscopy", *Proc. of ICATT'11 Conference*, Kyiv, Ukraine, 2011

OTHER CONFERENCE CONTRIBUTIONS

1) O. Sushko, R. Dubrovka, R. S. Donnan, "Hydration dynamics of proteins in solutions studied in 220-325 GHz band" The Physics of Soft and Biological Matter, Cambridge, UK, 14-16 April 2014, poster presentation;
2) O. Sushko, R. Dubrovka, R. S. Donnan, "Investigation of hydration shell thickness and electromagnetic field absorbance of solvated protein molecules in

the sub-THz frequency domain" IET Colloquium on Millimetre-wave and Terahertz Engineering & Technology, Liverpool, UK, 5th March 2014, oral presentation;

3) R. S. Donnan, B. Yang, A. McIntosh, O. Sushko, J. Qui, "Towards Developing a Dielectric Metrology for the Life Sciences" *PIERS 2013*, Stockholm, Sweden – oral presentation.
4) B. Yang, A. McIntosh, O. Sushko, R. S. Donnan, "Millimeter Wave and THz Metrology in Material, Chemical and Biological Sciences" *Biophotonics conference*, Edinburgh, UK, 2013, oral presentation;
5) O. Sushko, R. Dubrovka, R. Pickersgill, J. Qui, R. S. Donnan, "Potential of THz Radiation for Examination of Dissolved Proteins" *Today's RF Tomorrow's Medicine*, London, UK, 2013, poster presentation;
6) O. Sushko, R. Dubrovka, R. S. Donnan, "Deeper Insight into Material Optical Properties Extraction Procedure from THz-TDS in a Transmission Mode" *Photon 2012*, Durham, UK, oral presentation;
7) O. Sushko, R. Dubrovka, R. S. Donnan, "Lagrangian Dynamics Analysis of THz/far i.r. Spectra of Amino Acids and Peptides", *Dielectrics 2011*, Canterbury, UK, poster presentation;

Abstract

Terahertz frequency domain constitutes the least explored part of electromagnetic spectrum. At the same time plenty of physical phenomena occurs on picoseconds to nanosecond time-scale and have and can be monitored/controlled/studied by THz and sub-THz waves. Since the advent of photo-conductive generation followed by invention of the first THz-TDS system, research in this field made a huge progress, although still possess a considerable potential for growth. Alongside advances in generation and detection of THz radiation simulation tools are becoming increasingly important and facilitate interpretation of the experimental results.

Thesis comprises three related subjects, namely the processing of THz-TDS raw data, analysis of protein solvation dynamics by simulations and experimental investigation of water-protein solution at different concentrations. Experimental works in this thesis is performed using THz-TDS (normally covers 0.1-4 THz domain) and quasi-optical bench which covers the 75-325 GHz frequency bands. Molecular dynamics simulations were conducted in Gromacs package with a purely mechanical force field.

The thesis is organized in the following way: chapter 1 introduces THz frequency domain to the reader, by describing its location in the electromagnetic spectrum, the physical phenomena that falls to THz domain, the main applications of THz radiation and overview of the mechanism of interaction between THz waves and bio-molecules. Second chapter outlines the principles of operation, physical processes and areas of application of THz-TDS. It is completed with a detailed description of the THz-TDS available in our laboratory. Third chapter gives a general picture of data processing related to material parameter extraction from time-domain response of the sample recorded by THz-TDS. Then it goes into details of associated error analysis, introducing the uncertainty caused by utilization of approximated transfer function. The application of the accurate algorithm for sample thickness determination based on its THz response is also presented in the third chapter. The fourth chapter discusses the application of Gromacs molecular dynamics simulations for the study of solvation dynamics of four selected proteins, namely TRP-tail, TRP-cage, BPTI and lysozyme proteins. All the water molecules solvating protein are divided into buried in the protein interior structure and the ‘on-surface’ water molecules. The later is shown to have similar properties for all proteins, while the former serve as the origin for the differences in solvation dynamics of proteins. Further in this chapter the radius of hydration shell and its

dependence on the protein structure is investigated using vibrational density of states of solvating water molecules. The experimental investigation of the lysozyme, myoglobin and BSA proteins solutions performed over 0.22-0.325 THz domain using the PNA-driven quasi-optical bench is described in chapter 5. The relative absorption of protein molecules in solution and the hydration shell depth is also estimated. The last chapter concludes the thesis and outlines some future prospects.

Acknowledgements

Completing this dissertation would not have been possible without the help and assistance of many people. First of all I would like to mention my former supervisor Prof. Fedir Dubrovka who encouraged me to do PhD abroad and introduced me to research in the early days of my undergraduate studies. He has also kept continuous interest in my progress and often gives valuable advice.

Many thanks, of course, to my primary research supervisor, Dr. Robert Donnan, for his constant encouragement and support of any ideas and initiatives, stimulating me to become an independent researcher. Also, thanks for all the time spent correcting my English in the papers and thesis. Rob has always been timely in his replies to any requests I have had, keeping positive and friendly at all our meetings.

Naturally, I would also like to thank to my second supervisor, Dr. Rostyslav Dubrovka, for finding the time for discussions and critical reading of my papers; also supporting my research with his strong background in electromagnetics and antenna theory.

The head of our group and my independent assessor Prof. Clive Parini have thoroughly proof-read my intermediate reports and always given a critical, 'out-of-the-box' guidance that has been of much use.

I greatly appreciate the assistance of our lab manager Dr. Massimo Candotti and Mr. Tony Stone from workshop for their organized approach and always doing their best to meet any inquiry about equipment or time-slots.

I would also like to thank Kastriot Shala, senior PhD student in our group, for patiently explaining to me the basics of THz-TDS and for his advice and many fruitful discussions related to the writing of my first journal paper.

I am very obliged to Dr. Bin Yang and Dr. Wenfeng Sun whose fundamental knowledge in alignment of THz-TDS and quasi-optical bench was of significant importance for my experimental studies. Dr. Bin Yang, in particular, shared a lot of underlying information on the principles of THz generation and detection, protein hydration, data processing techniques, etc.

Thanks to Mr. Yang Zeng for being available when we needed to spend days realigning the THz-TDS optics. A separate thanks to Ms. Junyi Qiu who consistently provided me

with bio-related materials starting from protein samples to pipettes, and even Chinese sweets sometimes.

And of course many thanks to all my colleagues and friends, too many to mention, who made my time being a PhD student a more enjoyable and entertaining endeavor with many coffee/lunch-breaks, dinners and parties.

I would also like to acknowledge Dr. Mira Naftaly from the National Physical Laboratory for assisting in THz measurements on their site and providing insights into data processing and alignment of the THz spectrometer.

My warmest thanks and appreciation goes to my wife and parents for constantly supporting, inspiring and motivating me during my PhD studies.

Finally, I would like to acknowledge the Scholarship from the College Doctoral Training fund provided by Queen Mary University of London (now – Principal's studentship), which made my 3-year PhD program possible. Also, I have relied on the Antenna Group travel funds and Queen Mary Postgraduate Research funds that have supported me in traveling to many conferences and meetings in UK and abroad.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ACF – auto-correlation function

AFGP – antifreeze glycoprotein

BPTI – bovine pancreatic trypsin inhibitor protein

BSA – bovine serum albumin protein

CW – continuous wave

DFT – density functional theory

DNA – deoxyribonucleic acid

DR – dynamic range

EM – electro-magnetic

EO – electro-optic

FFT – fast Fourier transform

FIR – far infrared

FP – Fabry-Perot

FT – Fourier transform

H-bond – hydrogen bond

IR – infra red

$LiNbO_3$ – lithium niobate

LT-GaAs – low temperature grown gallium arsenide

MD – molecular dynamics

meV – millielectron-Volts;

MSD – mean square displacement

NMA – normal modes analysis

NMR – nuclear magnetic resonance spectroscopy

NTP – normal temperature and pressure

OF – objective function

PC – personal computer

PCA – photoconductive antenna

PDB – protein data bank

PE – polyethylene

PTFE – Polytetrafluoroethylene

QO – quasi-optics

SNR – signal-to-noise ratio

TD – time domain

TF – transfer function

THz – terahertz

THz-TDS – terahertz time-domain spectrometry

TPX – poly-4-methyl pentene-1

TRP – tryptophan protein

TV – total variation

VACF – velocity auto-correlation function

VDOS – vibrational density of states

VNA – vector network analyser

ZnTe – Zinc telluride

Chapter 1. Terahertz radiation

1.1 THz range of electromagnetic spectrum

The first occurrence of the term terahertz (THz) was roughly in the middle 1970s [Fleming'74], where THz was used to describe the spectral lines between 0.33 – 1.4 THz of gases such as H_2O, N_2O and SO_2. This designation has been also used by Ashley and Palka [Ashley'73] to address the resonant frequency of water lasers. However, spectroscopy in the far-infrared (IR) has been attempted since the early 1950s, driven by advances in Fourier-transform IR spectroscopy as described in a review by Loewenstein [Loewenstein'66]. Progress in THz technology has been largely limited mostly due to high atmospheric propagation losses and lack of efficient generation techniques [Kemp'06]. During the early stages THz technology found primary application in space sciences, where rich information of thermal emission lines of many light-weight molecules can be mapped by heterodyne and Fourier transform techniques [Siegel'02]. With the advent of ultrafast femtosecond lasers in the 1980s, THz spectroscopy gained the new functionality of coherent detection. This technique, called THz time-domain spectroscopy (THz-TDS), was based on optical excitation of voltage-biased semiconductors and was first demonstrated by Grischkowsky et al. [Exeter'89]. The progress and development of THz technology have been broadly described in a number of review papers [Siegel'02, Siegel'04, Tonochi'07] showing that nowadays this multi-disciplinary area is expanding its applications rapidly.

Today the term THz is used to describe the rather broad spectral domain between microwave and IR with a commonly used, but not strict definition, from 0.1 – 30 THz. An illustration of the THz frequency domain within the electromagnetic spectrum is shown in Fig. 1-1.

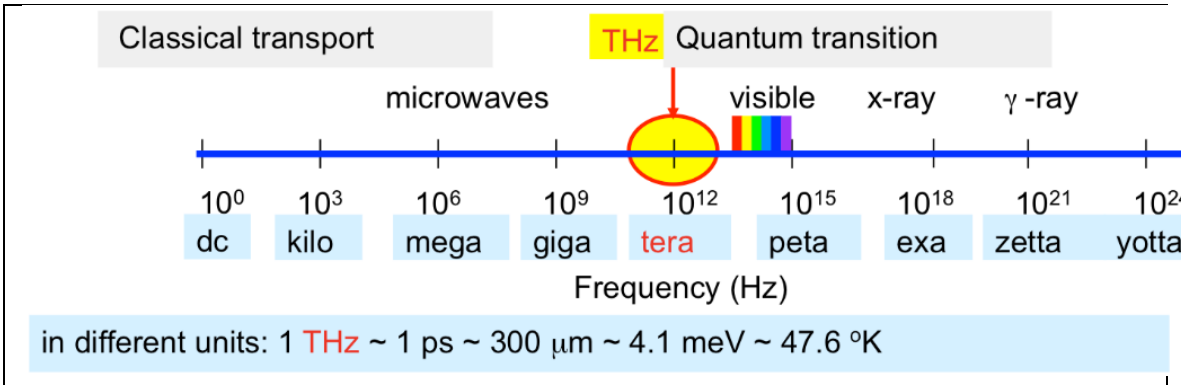


Figure 1-1. A brief illustration of the THz frequency domain [http://thz.phys.rpi.edu/images/thz_gap.png]

THz waves have many attractive properties and have been receiving increasing interest since the 1980s. A number of physical and chemical processes with lifetimes of the order of picoseconds can be probed by THz radiation. These include carrier-lifetime in semiconductors [Baxter'09], hydration dynamics of bio-molecules [Leitner'08], function-relevant motions in proteins [Cao'04], hydrogen bond fluctuations [Chakraborty'07], intermolecular vibrations in crystalline solids [Jin'10] and transient molecular dipole moments [Beard'02], etc. (Fig.1-2, 1-3).

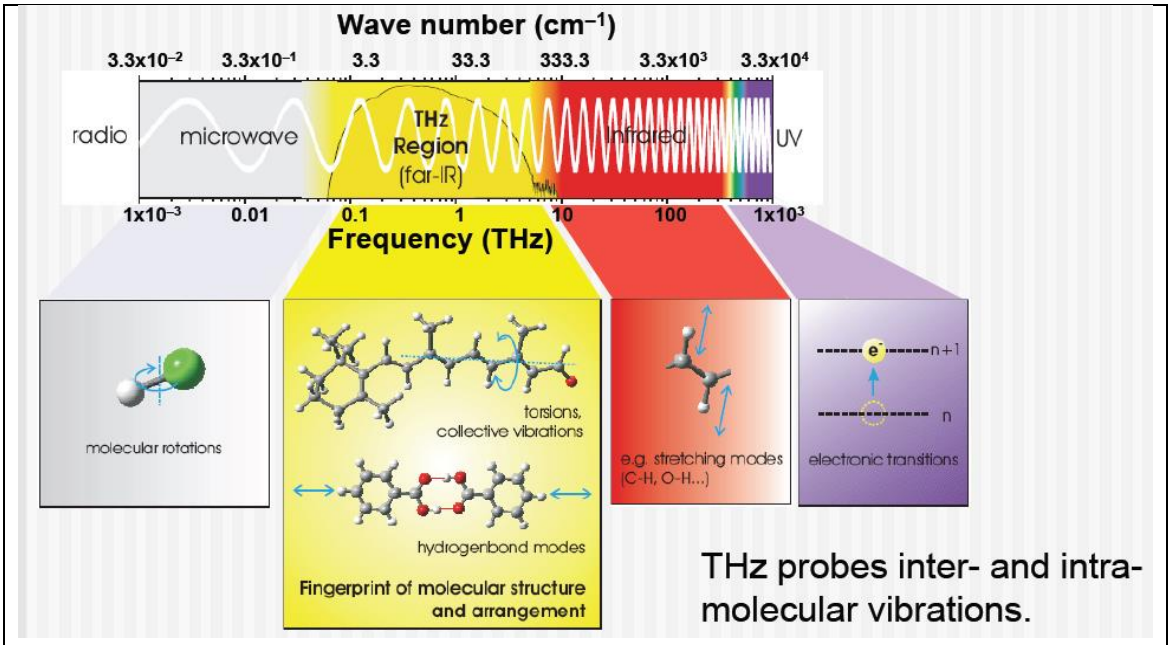


Figure 1-2. Molecular processes in different frequency domains [Davies'11]. THz/far-IR band spans the collective vibrations of bio-molecules that are responsible for their functions; also the vibrations of hydrogen-bonded molecules that represents lattice vibrations and drive solvation dynamics.

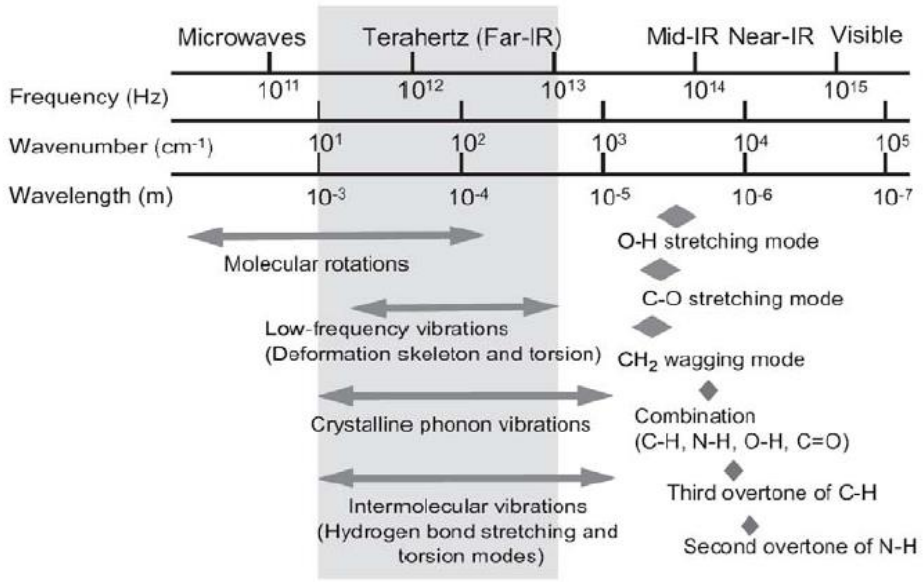


Figure 1-3. Characteristic infrared absorption frequencies of chemical compounds [Ueno'08].

Photon energy in the THz spectral domain is much lower than that of X-Rays, and so does not cause ionization effects in biological systems. This low (meV) energy level gives THz waves the advantage of non-destructive inspection of fragile biological samples such as protein and DNA. THz waves can also penetrate through a wide variety of materials without causing damage, enabling circuitry testing, explosive inspection, luggage and concealed weapons detection at airports. Short THz wavelength (0.1 – 1 mm) allows for a high resolution imaging [Federici'05]. Special techniques for THz near-field microscopy can reach 14 μm resolution at 0.7 THz [Blanchard'11] and even down to extreme sub-wavelength resolution of 150 nm [Kersting'08] achieved with few-cycle THz pulses of 3 THz bandwidth. Table 1-1 contains a summary on the wide applications of THz technologies that take advantage of the useful properties of THz radiation.

Table 1-1. Applications of THz Technologies [Jepsen'11].

Application Fields	**Tasks and Challenges**
Bio/Medical Medicine Infrastructure	Bio molecular, on-site THz imaging system, cancer diagnosis, medical services.
Security	Inspection system from hazardous materials/prohibited drugs, THz sensors and cameras.
Basic Science, Astronomy	Functional THz-TDS/Imaging systems, bio-molecular structure analysis, extraction of material properties in THz range, analysis of the space radiation at THz energies
Information Communication	Tb/s communication, THz wireless communication, THz sensor network, satellite communication, electromagnetic compatibility.
Industry Standards	THz-TDS/imaging system for industry applications, evaluation systems for nanomaterials, systems for controlling the quality of the product

Despite its many useful properties, the user of THz radiation must confront the major disadvantage it suffers – atmospheric absorption [Armstrong'12]. The attenuation

of electromagnetic (EM) waves at THz frequencies by the atmosphere under normal temperature and pressure (NTP) conditions is plotted at Fig. 1-4. Atmospheric attenuation is particularly high beyond 1 THz, reaching 10 dB/m.

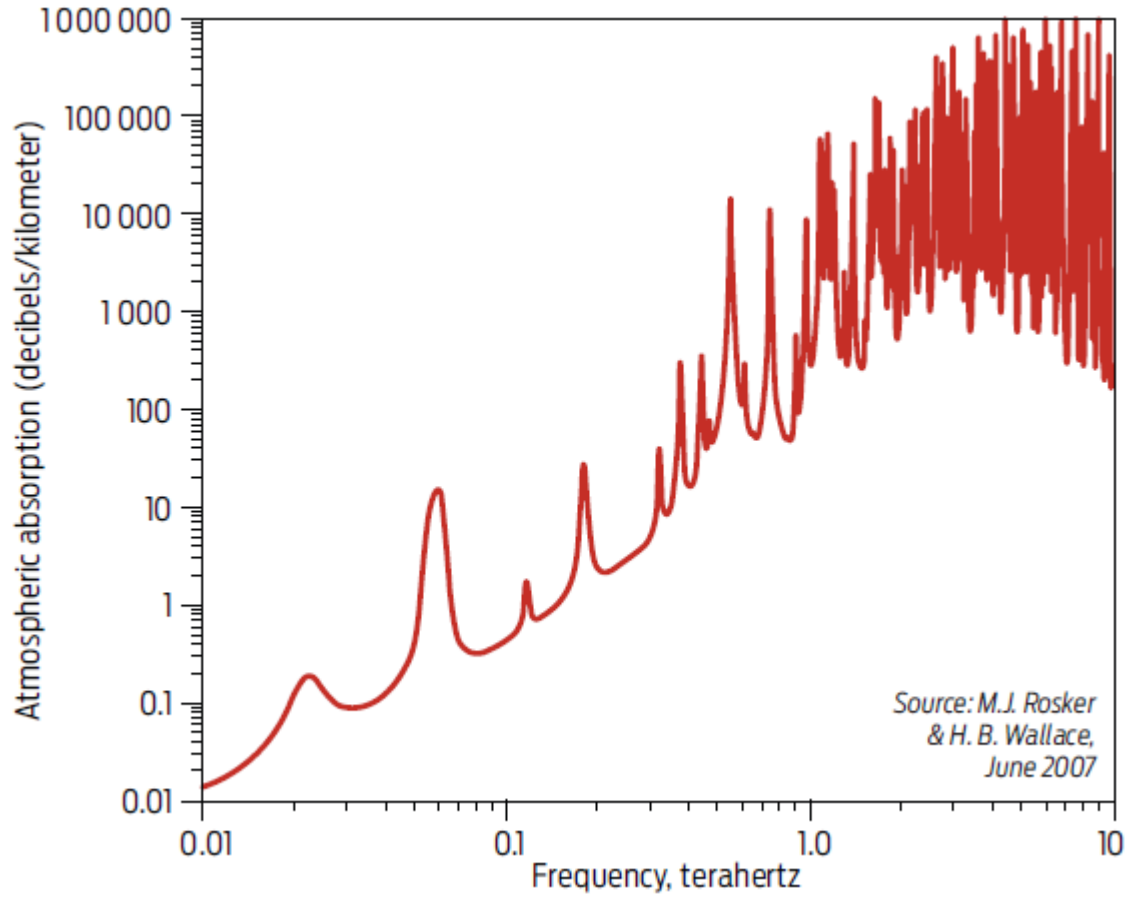


Figure 1-4. Atmospheric attenuation of THz radiation at ground level.

1.2 Application of THz technologies in biochemistry

THz waves have been applied to study a wide range of bio-samples covering everything from amino acids through cells and up to tissue. THz energies are being increasingly widely applied in bio-molecular studies. They are tuned to dipolar processes associated with hydration, molecular binding, temperature and conformational action. The progress in this field has been comprehensively reviewed by Siegel [Siegel'04], Markelz [Markelz'08] and more recently by Jepsen et al. [Jepsen'11]. As mentioned earlier THz response also clearly tracks inter- and intra-molecular vibrations of a small bio-molecules in the dry state. For instance, THz study of amino acids can provide information on chirality and crystallization [Yamaguchi'05]. THz absorption spectra of moderate to large sized bio-molecules (peptides containing 5 and more amino acids), do not resolve any features due to their dense overlapping and temperature

broadening. Instead, a continuously increasing THz absorption of protein powders has been observed and can be interpreted by a mode density function, obtained from classical MD simulations [Cao'04]. Cooling the sample to cryogenic temperatures can help to reveal sharp absorption signatures of molecular vibrations. The resonance frequencies of torsions, collective vibrations, H-bond vibrational modes that lie in the THz spectral domain represent a medium for sensing molecular structure and arrangement. This is potentially useful for identification of specific chemical components, in particular, in the detection of explosives and illicit drugs and for quality-control of pharmaceutical products. This information has been used in semi-empirical force fields and to refine the molecular mechanical force field that governs the atomic motion in molecular dynamics (MD) simulators. MD simulations often accompany experimental studies providing means for the assigning of absorption peaks, close analysis of hydration effects, folding events and interactions between the molecules.

Temperature-dependent studies have shown that proteins undergo a dynamical transition at around 220 K [He'08], the so-called glass temperature. Above this temperature proteins are dynamically activated and exhibit an increased atomic mobility. Ding et al. have determined the size of the solvation shell upon freezing peptide solutions by liquid nitrogen [Ding'10]. They rely on the fact that water molecules next to peptides have a suppressed ice-forming ability; therefore, when bulk water freezes, the water in a solvation shell still remains liquid.

The sensitivity of the THz radiation to the conformational state of a protein was demonstrated by Castro-Camus et al. [Castro-Camus'08]. The authors observed a distinct increase in absorption over 0.25 – 2 THz in photoactive yellow protein after illumination with blue light at 450 nm, which stimulates the protein to partly unfold. Higher absorption was then related to an increase in the density of delocalized vibrational modes in the partially unfolded state computed by normal mode analysis. Heyden et al. has investigated the 2.4 THz absorption of native and denatured ubiquitin and λ-repressor proteins by a p-Germanium laser [Heyden'10a] (achieved by varying pH of a solution). For both, native conformations clearly exhibited a higher THz response attributed to the modified water dynamics in the hydration shell of the unfolded state.

Apart from the crystalline or condensed state of bio-molecules, where they normally exhibit absorption features (dependent on molecule size), great effort has been made to study bio-molecules in their natural environment – i.e. dissolved. In this case all the low frequency intermolecular vibrations vanish (as molecules are not bounded to

each other in solution), while intra-molecular vibrational modes are masked by a strong water absorption in THz spectral domain (monotonically increasing from 90 cm^{-1} at 0.1 THz to 600 cm^{-1} at 3 THz). For this reason the layers of the solutions investigated are often reduced to a thickness of about 100 μm. It was established by many that hydration water around bio-molecules has different properties in comparison to bulk water. Based on this fact the solutions can be treated as three-component models consisting of: protein, bulk water and hydration water [Leitner'08]. Some studies [Heyden'08] have used this approach to determine the size and absorption of the hydration shell; others to extract the molar absorption of a protein [Zhang'06]. Niehus at al. [Niehues'11] have measured a concentration-dependent THz absorption of amino acids in solution. A firm correlation was found between THz response and such properties of a solute as polarity and hydrophobicity. The authors believe that THz spectroscopy (2.1 – 2.7 THz) can serve as a simple non-invasive test of hydrophobicity with no need for introducing additional markers. THz-TDS was also applied to detection of bio-molecules in solution by Arora et al. [Arora'12]. They have demonstrated quantitative detection of DNA with 697 and 133 base-pairs with a minimum detectable volume of 0.1 ng/μl. High sensitivity to hydration and water content results in potential applications of THz spectroscopy for detection of skin cancer and tooth imaging [Woodward'03]. The challenge of creating the compact THz imaging system suitable for the above applications still remains. The cross-checking of THz-TDS results with adopted orthogonal techniques is required, but often omitted [Falconer'12].

1.3 Interactions between THz Waves and Bio-molecules

Biological molecules interact with radiation of different frequencies with characteristic differences. Certain energies in the visible and ultraviolet region can cause electrons to be excited to higher energy orbitals. If the energy of a photon is sufficient, e.g. X-rays, molecule may be dissociated or ionized. Ionization is particularly harmful for organic molecules, since it creates chemically-active radicals, which can result in damage to other molecules. The most significant property of THz waves, with regard to interaction with bio-molecules, is that THz photons, being comparatively less energetic, (of the order of meV) are unable to cause ionization of a molecule. Photons in the THz/far infrared region carry less energy than that in the visible or ultraviolet region and can only excite vibrations in molecules. Microwave radiation is even less energetic and consequently can only cause rotations of molecules.

While new and brighter THz sources are emerging, the need is growing for detailed studies on the influence of THz electromagnetic radiation upon bio-molecules and living cells. This interaction is determined by the parameters of a THz beam (frequency, size, power, exposure time), and the properties of the exposed sample (refractive index, composition, absorption and scattering properties). The general mechanisms of interaction can be divided into absorption, scattering, thermal effects and so-called 'micro-thermal' or resonance influences [Wilmink'11]. The penetration-depth has often to be considered, which is approximately 100s of micrometers at lower THz frequencies and 10s of micrometers at higher THz frequencies (2-3 THz).

It is generally accepted that THz absorption by materials is an indication of the fluctuations of the system dipole moment [Heyden'10b]. In solution the rotational relaxation time of an average-sized protein is of the order of nanoseconds, in the dry state it is even slower. A THz wave has a period of oscillation of 1 picosecond. This is three orders of magnitude faster than the typical relaxation time of a protein. Therefore, as outlined by Heyden et al. [Heyden'12a], the alignment of its dipole moment lags the phase of the driving field and each individual dipole moment of a protein is effectively made stationary at THz frequencies. Most of solute-induced absorption changes, apart from the expulsion of water, are caused by the interfacial water (the first few water layers immediately adjacent to protein surface that have properties different from bulk water); also the contribution of protein electronic polarizabily is significant [Zhang'06]. In terms of absorption, the interaction of protein with THz radiation is realized mainly through the dipoles of separate internal group-vibrations, which densely populate the THz frequency domain. Special techniques like normal modes analysis (NMA) or principal component analysis (PCA) allow calculating the dipole derivatives and to estimate the absorption contribution of each internal vibration [He'08].

Once absorbed by a given medium, the electromagnetic energy is partially converted to heat, assuming no photochemical processes are involved. Changes in the temperature of a sample can in turn introduce uncertainties to the probed sample properties. Since most bio-tissues contain a high percentage of water, and considering that water is a dominant absorber at THz frequencies with respect to the bio-molecules themselves, it is useful to estimate the heating effect of THz radiation on water. Kristensen et al. have applied Kirchhoff's heat equation to model the influence of a terahertz beam on a sample of water [Torben'10]. The authors have shown that for a focused continuous wave (CW) beam of 0.5 mm diameter, the steady-state temperature increase is 1.8 K/mW, and that the frequency has a relatively large effect with 1.4

K/mW at 0.1 THz and 2 K/mW at 10 THz. They also discuss the application of their model to THz-TDS. Microwatts of pulsed THz power (typical for conventional THz-TDS setups), was argued to introduce at most, a negligible temperature rise of 1.8 mK, and only after the temperature reaches a steady-state value. Another approach of estimating the thermal effect of THz radiation has been suggested by Wilmink and Grundt [Wilmink'11]. According to them the temperature increment is proportional to irradiance, tissue absorption and exposure time and inversely proportional to the tissue density and specific heat capacity.

Along with absorption, scattering also weakens the THz response of the sample. Incident photons are scattered most strongly when the wavelength of the incident electromagnetic wave is comparable with the particle size. Proteins are normally much smaller than THz wavelengths. For this reason scattering often can be ignored in solutions of bio-molecules, where the response is dominated by highly-absorbing water. However for dry-state samples it is non-negligible. Mixing the powder of bio-molecules with polyethylene before forming pellets is a common technique in the characterization of dry-state chemicals. Under such circumstances scattering in the sample originates from mismatch of the refractive index among constituents of the sample. Therefore, it is usually well pronounced for inhomogeneous samples where two materials with different properties are mixed or where voids are present in the sample. Scattering is the reason for asymmetric peaks profile in THz absorption curves as discussed by Franz et al. [Franz'08]. Based on information of frequency, particle size and the contrast in refractive index of mixed materials, the authors proposed a procedure for eliminating the scattering component from absorption spectra. Kaushik et al. suggested an iterative multilevel wavelet transformation for estimation of scattering baseline from THz-TDS measurements [Kaushik'12a]. Alternatively, a practical method for mitigating the scattering effects without the need for any *a priori* information about the material properties was introduced [Kaushik'12b]. This method assumes that the material under test has sharp features in THz domain. Scattering effects can be minimized in several ways: by using finely-milled samples ensuring that powders are well mixed to maximize homogeneous distribution of absorbing particles; compressing powders to form compact pellets, thus reducing porosity; and averaging over several measurements taken for different sample positions [Shen'08].

In the early 1970s Frolich raised the question of the possibility of non-ionizing, THz/sub-mm waves being employed to interfere with bio-molecules in a destructive but non-thermal manner [Frohlich'75]. This interaction has been referred as 'micro-thermal'

or ‘resonance’ and is little studied due to lack of experimental apparatus to probe electromagnetic fields and dynamics at the molecular level. It is stimulated by the experimental evidence on a macroscopic level that cannot be readily interpreted only in terms of temperature increase [Swanson’11, Bock’10]. This “non-thermal” interaction is thought to result from resonant phenomena driven by external EM fields. Several recent studies have aimed at developing a theoretical background for these effects on DNA [Swanson’11, Alexandrov’10]. They confirm, with some caveats, that under certain conditions THz radiation might be able to destabilize DNA breathing modes.

Chapter 2. THz Time Domain Spectroscopy

2.1 Principles of operation

The conventional setup of THz-TDS was first introduced by Grischkowsky and co-workers. Due to its coherent nature it provides information of complex refractive index of measured samples (that can be correspondingly converted to complex dielectric function and conductivity). Nowadays THz-TDS systems are more technologically mature, namely: laser beams are fiber-coupled directly to an emitter; faster scanning time; emitters integrated with special plasmonic grating to enhance THz output; lenses are integrated on both emitter and detector; the decrease of noise floor is achieved by cooling the emitter and new materials used, etc. However, the basic principles of operation remain the same. The main technological pursuit at present is to make THz-TDS a table-top, compact and more affordable instrument. A schematic illustration of the setup of a THz-TDS in transmission mode is shown in Fig. 2-2. The pulsed titanium-Sapphire laser produces femto-second pulses, which are then redirected into separate optical paths by a beam-splitter.

The pump-beam illuminates a biased photoconductive antenna (PCA) to induce free-carriers. Due to the fact that electrons are more mobile than holes, usually the laser is incident on the photoconductor near the anode. A bias voltage applied between two parallel-strip plate-electrodes accelerates the carriers to produce a transient current that is proportional to the time-derivative of the dipole moment induced in the emitter, i.e.:

$$\vec{I}(t) = \frac{\partial \vec{p}(t)}{\partial t}.$$

The bias voltage is modulated at a rate of a few kHz to enhance SNR and facilitate lock-in reading. Alternatively a mechanical chopper in the pump beam path can be used for the same purposes. According to basic antenna theory, accelerated current gives rise to the radiation of an electromagnetic wave with electric field governed by [Zhang'09]:

$$\vec{E}(t) = \frac{1}{4\pi\varepsilon_0 c^2 r^3}\vec{r} \times \left(\vec{r} \times \frac{\partial \vec{I}(t)}{\partial t}\right).$$

In some cases a dielectric lens is mounted on top of an antenna to confine most of the THz waves in a certain solid angle. The lens also enhances efficiency by providing better coupling between emitter-material and free-space.

An interesting analogy can be provided here to enhance understanding of origin of THz radiation. Effectively, the PCA can be treated as a small capacitor charged by a

bias voltage and accumulating energy equivalent to $0.5CV^2$. The capacitor is then discharged by a laser-induced conductivity and part of the energy is radiated; accordingly the THz energy originates from the power supply recharging the capacitor.

Off-axis parabolic mirrors are used to direct the THz beam and to focus radiation onto a sample. A sample is usually positioned at the focus of the second parabolic mirror (OM2), or, if the sample is extensive, in the collimated beam between OM3 and OM4. After propagating through the sample, the THz wave is again refocused by off-axis parabolic mirrors to an electro-optic (EO) ZnTe detecting crystal. The pump beam travels through a time-delay stage so that arrival of the THz radiation to the ZnTe is delayed by Δt relative to the probe beam. This allows a step-by-step recording of THz wave intensity (Fig. 2-3). THz radiation induces birefringence in the nonlinear ZnTe due to the Pockel effect [Chen'08], characteristic of EO crystals. As a result, a cross-polarization component will be produced in the ZnTe plate that changes the balance of 2-port photo-detector (without THz wave two orthogonal polarizations of laser beam are balanced) (Fig. 2-4). ZnTe is often used for EO detection due to the small group velocity mismatch between the THz wave and the laser pulse in the crystal (0.4 ps/mm (15 ps/mm for GaAs)). After propagating through a quarter wave plate, the EM wave acquires elliptical polarization. A birefringence crystal (Wollaston prism) resolves the two orthogonal wave components. Balanced photo-detectors measure the intensity difference between these (ordinary and extraordinary rays) of the probe pulse, which is proportional to the applied THz field amplitude, (Fig. 2-4) and can be evaluated as [Winnewisser'97]:

$$\Delta I = I_{probe} \cdot E_{THz} \cdot 2\pi\nu \cdot n^3 \cdot d \cdot r_x/2c,$$

where ν is the optical probe-beam frequency; n and d are respectively the refractive index and thickness of the detecting crystal; r_x is a component of electro-optic tensor.

As a result of a single scan, a discrete THz waveform is obtained due to step-like motion of a delay-stage. A discrete Fourier transform is then applied to the time-domain response to provide discrete amplitude and phase spectra. Alternatively, in some systems, a PCA can be used as a detector. In this case the receiving PCA is not biased and current is induced by an incident THz wave. Also, EO crystal can be used for THz generation instead of by a PCA.

For measurements of highly absorbing materials, or even metals, THz-TDS in reflection mode is used [Jeon'98]. The same optical circuit is employed as in transmission mode, but the THz beam is reflected from the sample instead of propagating through it. Attenuated total reflection (ATR) mode is another modification

of THz-TDS that can be integrated into the system in transmission mode [Nagai'06]. It utilizes a special triangular prism (Fig. 2-1) that focuses the THz beam onto the sample with the ensuing reflected beam propagating collinearly with initial beam. Appropriate extraction procedures for estimating material properties accompany each THz-TDS configuration.

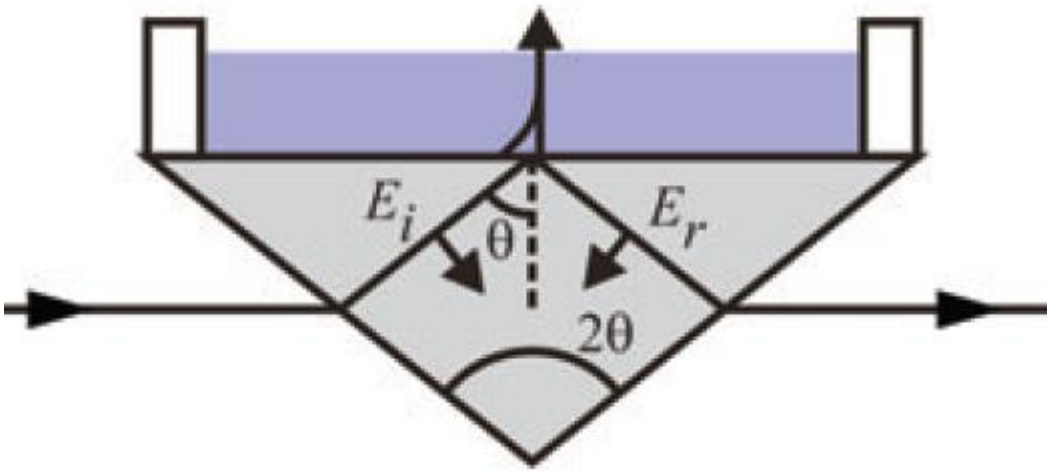


Figure 2-1. Schematic of the focusing prism used in attenuated total reflection spectroscopy (picture adopted from [Nagai'06])

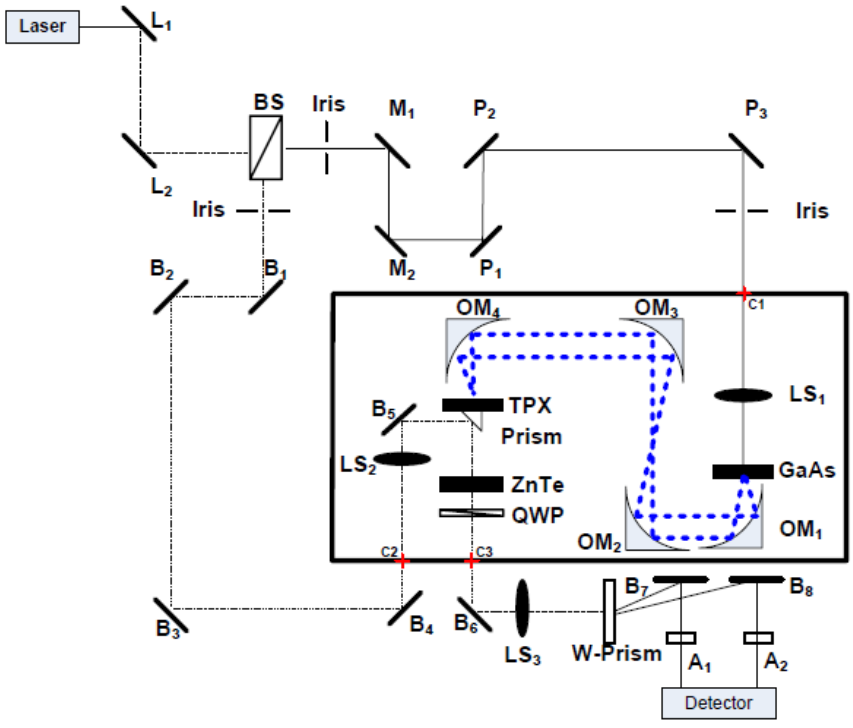


Figure 2-2. Schematic diagram of a THz-TDS system operating in transmission mode as configured in our THz laboratory. L, P, B stand for flat reflecting mirrors, M – flat mirrors of delay stage, LS – focusing lenses, A – attenuators, BS – beam splitter, QWP – quarter-wave plate, OM – off-axis parabolic mirrors. The THz beam path is marked in blue. An enclosing box for controlling atmosphere about the sample under test, is shown by the bold black line.

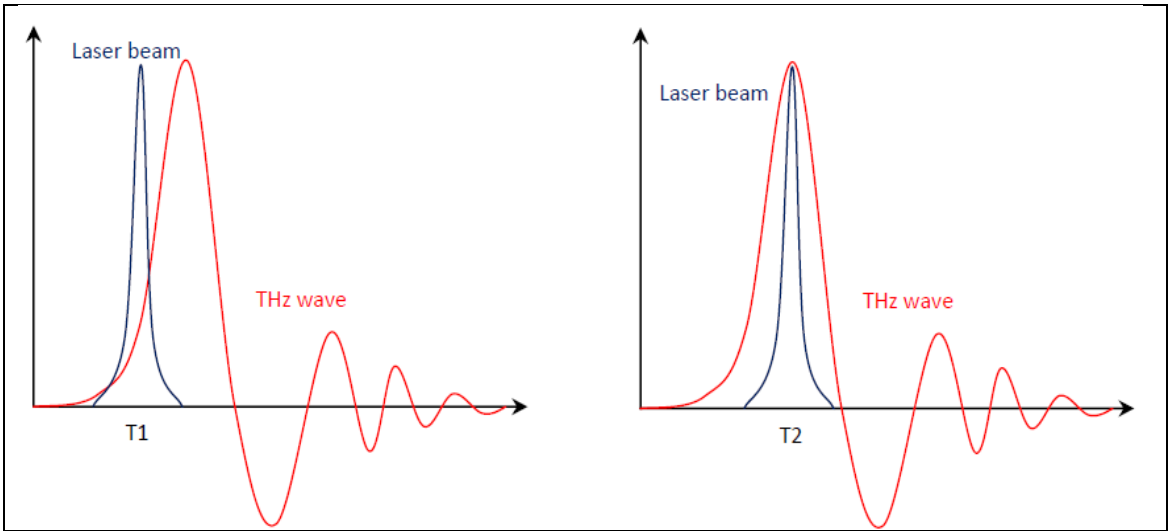


Figure 2-3. The overlap of THz wave and a probe beam at the ZnTe crystal at different positions of delay-stage. THz radiation induces birefringence of ZnTe that is probed by the laser beam.

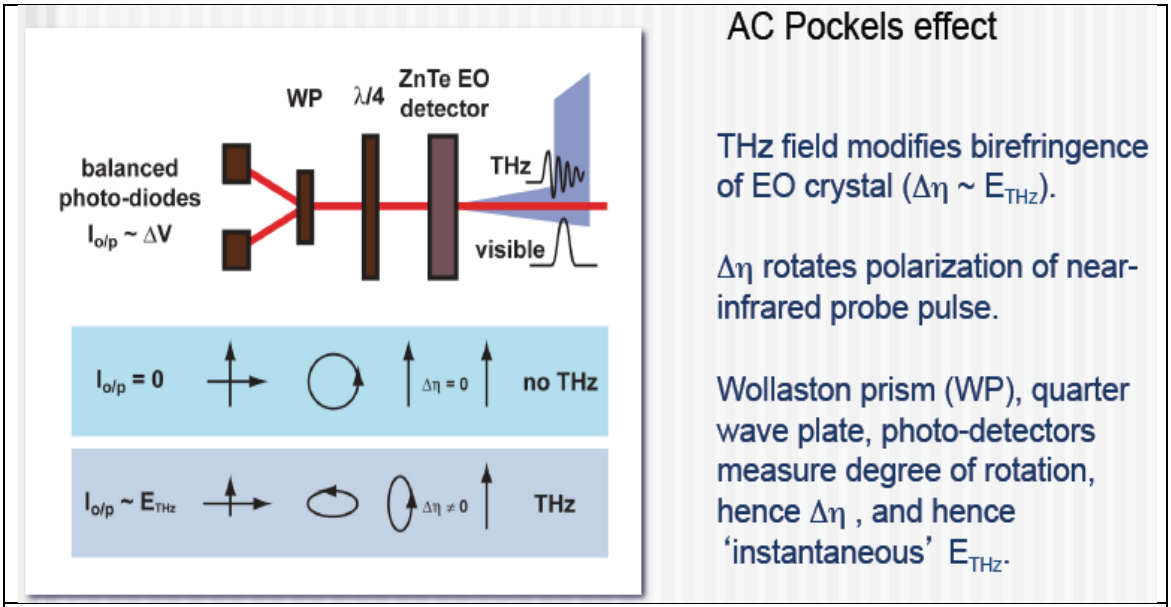


Figure 2-4. The detection of THz wave using electro-optic crystal [Davies'11].

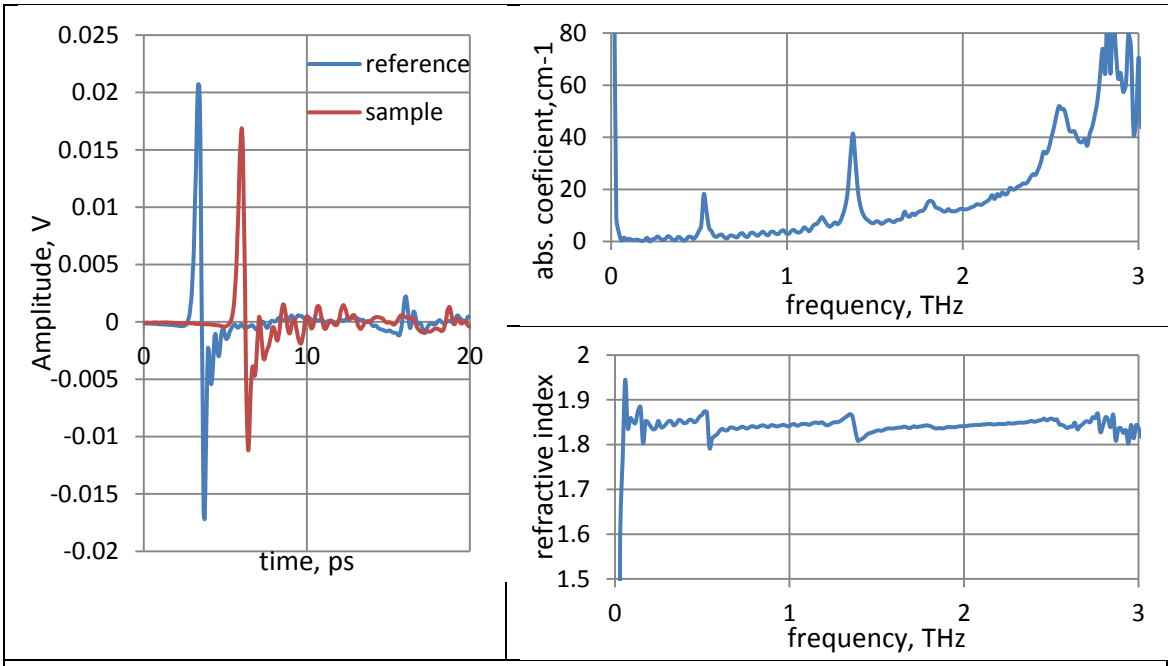


Figure 2-5. Left: time-domain response of an empty (free-path) system (blue) and for lactose pellet (red). Right: absorption coefficient and real part of refractive index of lactose extracted from the time-domain response.

In order to extract the spectral information of the sample, a reference measurement with the sample removed is first performed and a Fourier transform then applied to retrieve the spectra. The same procedure is performed with the sample in place. The ratio of these spectra reveals the interaction between electromagnetic waves with the sample in the THz spectral domain. Fig. 2-5 shows time-domain responses of an empty sample-holder (reference measurement) and lactose pellet and material dispersion parameters of lactose. The lactose absorption spectrum exhibits well-defined peaks corresponding to the intermolecular vibrations in the lactose crystals. Refractive index of lactose is on average 1.85 over 0.1 – 3 THz and has a phase jumps at the positions of the absorption peaks (since real and imaginary parts of refractive index are related via Kramers-Kronig formulae).

Again we emphasize a useful feature of THz-TDS is that it affords coherent measurement. This means a recorded response contains both amplitude and phase information of signal-beam interaction with a sample. These, in turn, allow extracting both real and imaginary parts of the refractive index.

2.2 Application of THz-TDS

During last decade THz-TDS has become a widely used technique for experimental investigation of bio-molecules, polymers and semiconductors. The main

interests of research groups regarding bio-molecules lies in this field of discrimination between conformational states of molecules [Ebbinghaus'08]; investigation of internal motions of molecules and assignment of absorption peaks to these [Ding'11]; study of the lowest frequency vibrational modes of molecules which are related to their function [Markelz'02, Kawaguchi'10]. Polymer characterization by THz-TDS is often used to provide accurate information on materials properties for applications in THz optics as sample-holders, beam-splitters, lenses, etc. [Naftaly'05, Jin'06]. Semiconductors are investigated mainly to characterize carrier mobility and emission properties [Shan'04]. In addition to the general applications of THz radiation described in Chapter 1, the most wide-spread and diverse THz-TDS applications are outlined below.

THz spectroscopy has high potential for detection and identification of intermolecular hydrogen bonds in unknown mixture samples (e.g. lactose α monohydrate, acetylsalicylic acid, sucrose and tartaric acid compressed pellets) [Ueno'08]. These studies showed that THz-TDS is a sensitive probe for detection of hydrogen bond vibrations that form crystalline or quasi-crystalline structures in such material. THz-TDS can also be used to characterize crystalline properties of drugs and excipients. Different polymorphic forms of a drug can be readily distinguished and quantified [Zeitler'07]. It was also shown that structural changes introduced by chemical reactions and molecule associations can strongly affect terahertz spectra, causing significant changes in absorption peak intensities and shifts in peak positions [Bykhovski'10].

THz-TDS studies are often complemented with corresponding molecular dynamics simulation to interpret spectra. Study by King et al. [King'11] demonstrates the capabilities of solid-state density functional theory (DFT) to assign observed THz absorption features and to uncover the underlying nature of the vibrational motions contributing to absorptions in the THz regime. The overall correlations of the simulated (*S*)-(+) ibuprofen (*RS*)-ibuprofen spectra with the experimental THz spectra are of high quality. Despite the apparent difficulties in the calculation of accurate relative intensities, all experimental peaks can be confidently assigned to calculated vibrational modes. Yunfen He et al. proved the possibility of investigating the presence of structural collective motions, as a function of oxidation and hydration, on a picosecond time scale for the proteins, using THz-TDS and molecular dynamics simulations [He'10].

Many bio-molecules have been studied in water or salt solutions environments [Choi'12, Yamamoto'12]. Based on frequency- and concentration-resolved THz

measurements, the information on bio-molecule molar absorption [Xu'06b], hydration shell size [Ebbinghaus'07] and of protein water interaction can be obtained. In [Xu'06a], molar extinction of solvated protein lysozyme (1001 atoms) between 0.075 and 3.72 THz (2.5-124 cm^{-1}), was measured and direct comparison to several published theoretical models based on molecular dynamics simulations and normal-mode analysis was made. The existence of dense, overlapping normal modes in the terahertz frequency domain was confirmed. The measured spectrum, while in rough, qualitative agreement with these models, differs in detail. The ability to discriminate between different functional states of proteins using spectra recorded by THz-TDS has been shown. Protein solutions at different pH ranging from 2 to 7 have been seen to exhibit distinct absorption due to the protein being in a folded or unfolded state [Ebbinghaus'08]. Separately, the unfolding event of a photo-active protein in solution has been stimulated by a diode-light and respective changes in the THz absorption detected [Castro-Camus'08].

An attempt to apply THz spectroscopy to identify the diseased tissue of brain grain matter is described in Gretel et. al. [Gretel'09]. A clear distinction between the absorption coefficients of diseased and healthy tissue is evident. However, it is not possible to conclude that this distinction is due to protein plaques in the diseased samples. Observed differences may be caused by tissue atrophy resulting in less dense sample for diseased tissue.

THz-TDS was also successfully used to study glasses, lubricating oils, and polymers. Relationships were observed between the composition and structure of the materials studied and their THz absorption spectra and refractive indices. THz-TDS was therefore shown to be a valuable tool in the study of materials [Naftaly'07].

Among other materials, semiconductors have been investigated extensively owing to their unique properties in THz frequency domain. Carrier mobility and density have been particularly studied [Exeter'90a,b]. THz-TDS spectrometry of thin films of semiconductors gives access to both thickness (due to Fabry-Perot reflections) and doping properties [Jeon'97].

Some groups have extended their conventional THz-TDS setup for imaging purposes by combining the system with 2D in-plane motor stages [Federici'05, Zeitler'07]. This is used for hydration measurements, identification of hidden object, investigation of pellet coatings and differentiation between different powders, etc.

In this thesis, the main focus of experimental studies is solutions of bio-molecules and semiconductors.

2.3 Measurement facilities in the QML laboratory

For experimental measurement we use the THz-TDS installed at QML with the assistance of National Physics Laboratory (NPL) scientists. A photo of it is shown in Fig. 2-6. The schematic diagram of the system was shown earlier in Fig. 2-2. Primarily the system as operated here is in transmission mode, whilst additional parabolic mirrors and flat-reflectors (not shown in the diagram) are installed to operate reflection-mode measurements if necessary. This complementary extension is used to characterize highly absorbing samples like doped semiconductors, metal films, polar liquids, etc. The main features and characteristics of this spectrometer are:

- Typical operating frequency domain: 0.1 – 4.0 THz;
- Maximum dynamic range (DR) is 25-30 dB (based on field spectral amplitude), SNR is normally around 25 dB.
- Typical resolution is 14 GHz (scan size 10.24 mm, step 10 μm), maximum achievable is 1-2 GHz;
- Laser source – class 4 Ti:Sapphire femtosecond pulsed laser with adjustable wavelength in the range 750-850 nm; pulse repetition rate is 80 MHz; average power is about 1 W;
- Motorized delay stage – maximum traveled distance is 15 cm.
- THz emitter – biased LT-GaAs photoconductive antenna, biased voltage typically 200V, 0.5 mm thickness. The gap size is approximately 0.5 mm which makes the laser beam positioning easier, compared to common 20-50 μm gaps.
- ZnTe (2 mm thickness, (110) crystallographic orientation) crystal is used as THz electro-optic detector. Such a thickness is chosen to allow enough interaction length of probe beam and THz wave in the crystal.
- The probe beam is directed to the ZnTe crystal using a TPX plate with a small reflecting prism on top of it (see Fig. 2-2). THz radiation can easily penetrate through TPX plate due to its extremely low absorption coefficient in the THz frequency domain.
- Overall, the system includes 17 flat mirrors, 5 parabolic mirrors and 3 lenses in the current setup;
- The THz table is floated/air-cushioned to avoid mechanical vibration from the floor during measurements;
- A THz box envelops experiments and can be purged with nitrogen or dry air to mitigate water vapor influence on the sample spectrum. Water vapor causes sharp absorption peaks in the THz frequency domain as shown at Fig. 1-4;

- A balanced Nirvana photo-receiver (model 2007) is utilized for recording of the time-domain response (the specifications of this photo-receiver can be found online via the link: http://assets.newport.com/webDocuments-EN/images/2007_And_2017_User_Manual_RevC.pdf);
- Data reading from the Nirvana detector is realized via a lock-in amplifier and is delivered to the PC through a LabVIEW interface. Special functionalities in LabVIEW provide means for signal averaging on points and scans.
- Material properties' extraction is performed in a Mathcad simulation package.

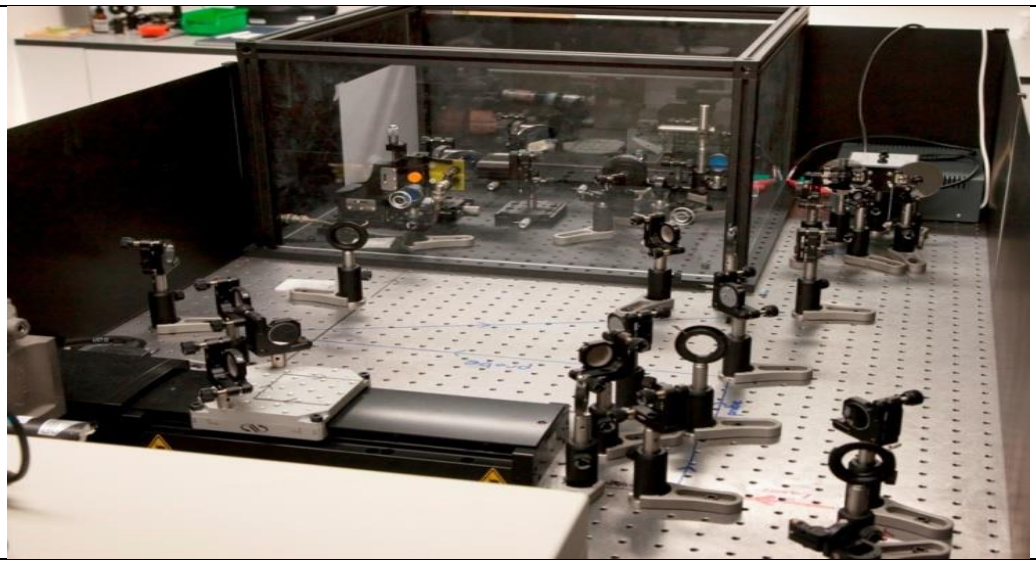


Figure 2-6. THz-TDS system in our laboratory.

Another experimental tool used to acquire THz/sub-THz properties of the materials is a quasi-optical (QO) bench driven by a vector network analyzer (VNA[1]). Figure 2-7 shows a photo of a QO bench in transmission geometry and its respective schematics. The VNA has a 10 MHz to 43.5 GHz operating frequency domain. Special millimetre-wave frequency extension heads (shown in blue in the picture), cover the following waveguide bands: 50-75 GHz; 75-110 GHz; 110-170 GHz and 220-325 GHz. Rectangular waveguide outputs of the extension-heads are terminated with high-gain corrugated horns. Two off-axis, ellipsoidal mirrors are used to focus source radiation at the sample location. A further two ellipsoidal mirrors direct the radiation to the receiving horn. In reflection mode only one extension head and the first two ellipsoidal mirrors are required. The system enables acquisition of a full set of complex S-parameters of the sample in the above frequency bands. Furthermore, with appropriate data processing techniques, the complex dielectric properties of the sample can be

[1] model HP N5244A

estimated. This setup offers low noise performance (typically 0.01 dB) and high dynamic range (>70 dB, depends on the frequency domain).

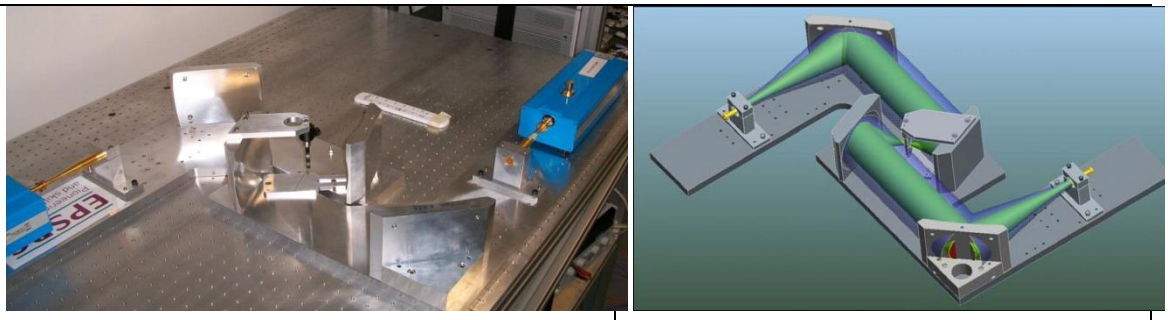


Figure 2-7. Left: Picture of the physical quasi-optical bench in a transmission mode and right: a schematic of the 98% power-level envelope of the fundamental Guassian beam-mode propagating through the circuit.

A note has to be made here with regard to the calculation of DR and SNR for both systems. A pertinent paper by M. Naftaly et. al. has been devoted to the topic of methodologies for determination of DR and SNR for THz-TDS [Naftaly'09]. First of all, these two parameters can be determined in both time and frequency domains. If the quantity of interest can be extracted directly from time-domain traces, then SNR and DR have to be determined from these data; and if the user utilizes spectroscopic information then SNR and DR have to be calculated from the FT amplitude spectrum. Since our measurements mainly concern frequency-domain information, the SNR and DR has been estimated as follows [Naftaly'09]:

$$SNR = \frac{mean\ FT\ amplitude}{standard\ deviation\ of\ amplitude}$$

$$DR = \frac{mean\ FT\ amplitude}{noise\ floor}$$

Since amplitude is dispersive, the SNR and DR were determined at maximum amplitude, i. e. around 1 THz. So the DR defines the maximum attenuation of the sample that can be characterized, and SNR – how accurately it can be done. On the other hand DR from VNA readings has the same nature as DR for TDS, but the RMS of the noise is used instead of SNR for TDS. Noise RMS values can be converted to SNR, as defined for TDS, but only for particular measurement and not for the system in general. For instance, for a 100 μm thick water layer the SNR at 0.3 THz is 25 dB and 55 dB for TDS and VNA respectively. Also VNA measurements have a superior DR by a few orders of magnitude (typically 70-80 dB in 0.22-0.325 THz band).

Several sample holders provide specialized handling for liquid and solid spectrometry. For liquids, a TPX (poly-4-methyl pentene-1) is used as a window material due its unique properties in THz spectral domain. Its absorption is very low, less than 1 cm^{-1} and index of refraction is 1.46 and both are non-dispersive in over 0.1 – 4 THz. Also TPX is easily mechanically-shaped (processed). The picture of the liquid sample holder purchased from Bruker is shown in Fig. 2-8. Several rotational and translational holders allow investigation of a set of rotation angles (360° in elevation and in-plane of the sample; +/-20° zenith angle measured from vertical), as required for emission ellipsometry of semiconductors.

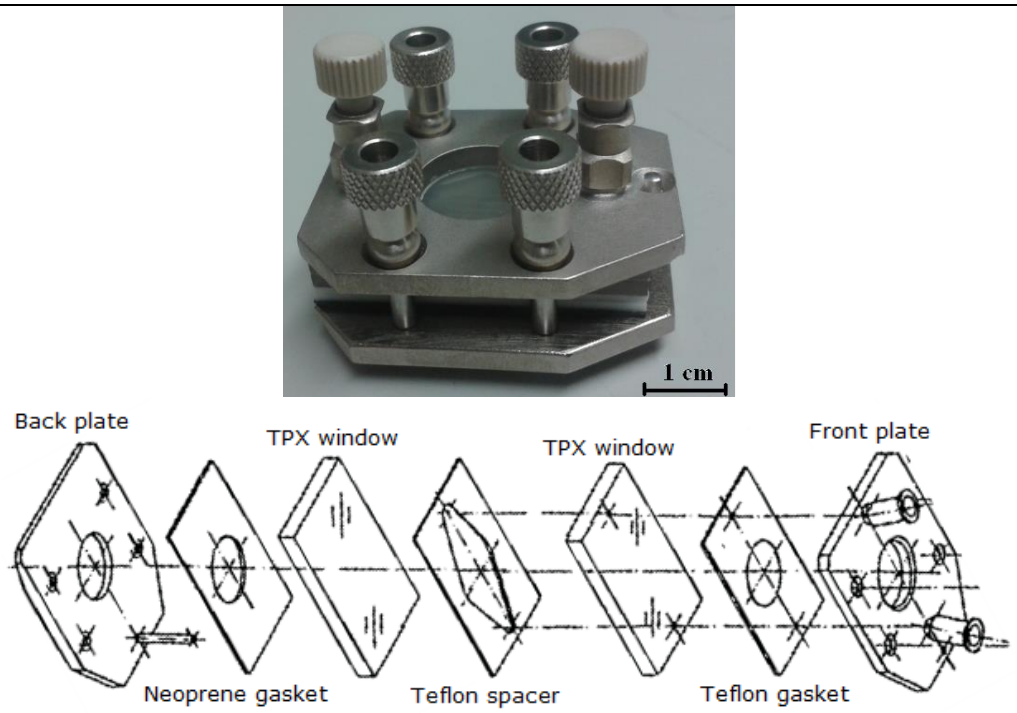


Figure 2-8. The picture and schematics of commercial liquid sample holder from Bruker company, PTFE spacers of different thickness are available, ranging from 25 μm to 1mm.

Chapter 3. Signal processing in relation to THz-TDS

The whole measurement procedure of material characterisation using THz-TDS is complex by virtue of sensitivity to attendant environmental conditions and optical alignment. The chief aim in most TDS applications then, is to obtain the complex refractive index of a given material in the most accurate manner possible over a broad frequency domain. For this, advanced signal processing techniques [Pupeza'07] and a thorough error analysis [Withayachumnankul'08a] are required. With appropriate data processing it is possible to: decrease signal noise; precisely identify sample-thickness numerically [Duvillaret'99, Duvillaret'00]; and acquire more accurate estimate of the complex refractive index.

Improving procedure and refining data-analysis are ongoing, as there is no generally accepted standard of optical constants determination. This chapter will focus on data analysis procedures in an attempt to perform a thorough analysis of time-domain (TD) response of a sample and maximize the accuracy of THz-TDS measurements overall. The dependence of final parameter estimates on initial guesses required for the iterative-fitting procedure is outlined. The instances when the usual algorithms fail to predict correct values of parameters due to poly-root behaviour of the TF are discussed, and a way is presented to overcome this drawback. Another issue addressed in this chapter is the undertaking of a comprehensive error analysis. Procedures are outlined to minimize/trace most sources of error. In particular, the uncertainty of final parameters, resulting from misinterpretation of the Fabry-Perot term in the TF, is pointed out and quantified for the first time.

3.1 Simple material parameters extraction procedure

The conventional approach for determination of the complex refractive index $\tilde{n}(\nu) = n(\nu) - ik(\nu)$ of a material is based on the ratio comparison of spectra of the sample with respect to a reference – which for transmission is served by the absence of a sample. This ratio constitutes a transfer function (TF):

$$\tilde{H}_{exp}(\nu) = \frac{\tilde{E}_s(\nu)}{\tilde{E}_r(\nu)}, \tag{3.1}$$

where $\tilde{E}_s(\nu)$ and $\tilde{E}_r(\nu)$ are the complex spectra of sample and reference, obtained by FT of respective TD responses. Considering that propagation of an electromagnetic wave through a medium is described by [Born'99]:

$$\tilde{A}(\nu) = exp\left(\frac{-i2\pi\nu\tilde{n}(\nu)d}{c}\right),$$

the analytical TF is augmented to be:

$$\widetilde{H}(\nu) = t_{12}(\nu)t_{21}(\nu)\exp\left(\frac{i2\pi\nu d(\tilde{n}(\nu) - n_{air})}{c}\right), \tag{3.2}$$

where $t_{12}(\nu)$ and $t_{21}(\nu)$ are the Fresnel transmission coefficients associated with the front and back boundary interfaces between sample and host medium to sample; c, is the usual free-space phase velocity of light; d, is the sample depth; $\tilde{n}(\nu)$ is the complex refractive index of the sample under test and n_{air} is a refractive index of air.

Determination of a sample's complex refractive index can be obtained by matching experimental and analytical TFs at corresponding discrete values of frequency. The resulting well-known formulae for refractive index n and extinction coefficient k are:

$$n(\nu) = 1 + \frac{c}{2\pi\nu d}\left(\varphi_s(\nu) - \varphi_r(\nu)\right) \tag{3.3}$$

$$k(\nu) = -\frac{c}{2\pi\nu d}\ln\left[\frac{\left|\tilde{E}_s(\nu)\right|}{\left|\tilde{E}_r(\nu)\right|}\frac{(1+n(\nu))^2}{4n(\nu)}\right], \tag{3.4}$$

where $\varphi_s(\nu)$ and $\varphi_r(\nu)$ are the unwrapped phase of a sample and reference spectra. Historically, the property of radiation absorption by a sample is characterized by an absorption coefficient α which is related to k by $\alpha = 4\pi\nu k/c$. Equations (3.3) & (3.4) are often used to obtain the optical properties of a material using THz-TDS [Jepsen'05, Naftaly'07a, Xie'12]. These expressions however do not take into account the multiple internal reflections within a sample, akin to the Fabry-Perot (FP) effect. These second and higher order internal reflections of the main beam pulse are simply ignored and various signal conditioning, such as zero-padding, signal windowing, etc. are applied to the time-domain response before evaluating of n and k. A further assumption here is that in equation (3.2) the Fresnel coefficients are taken as purely real. The relative uncertainty of this approach varies in the range of a few percent (shown below). Such accuracy is insufficient for analytical purposes in materials' characterization and identification or for deployment of such material in precision quasi-optical systems.

While eqn. (3.3) shows n to depend on the phase difference for a signal beam propagating through a sample-depth relative to an equivalent free-space depth, k is

more complex. It is itself a function of n and, importantly, of relative spectral amplitudes, that are prone to noise (see eqn. (3.4)). The extinction coefficient, therefore, is more sensitive to measurement procedure, instrument alignment, environmental conditions, etc. that have a direct effect on signal strength. Performing short-scan[2] measurements in an attempt to lessen the FP effect and/or applying signal processing techniques (e.g., zero padding, signal truncations, windowing), can lead to significant errors. This is especially accentuated when sharp absorption peaks are present such as those in lactose, amino acids, etc. (as shown below).

3.2 A procedure for more accurate determination of $\tilde{n}_s$

The way forward is to account for the presence of internal reflections and complex Fresnel coefficients. A more advanced analysis, therefore, is based on regression methods where FP-like effects are incorporated [Pupeza'07, Dorney'01]. In this case the modified TF is:

$$\tilde{H}_{th}(\nu) = \tilde{t}_{12}(\nu)\tilde{t}_{21}(\nu)exp\left(-\frac{i2\pi\nu d(\tilde{n}(\nu)-n_{air})}{c}\right) \cdot \sum_{l=0}^{m}\left[\tilde{r}^2(\nu)exp\left(-\frac{i4d\pi\nu\tilde{n}(\nu)}{c}\right)\right]^l, \quad (3.5)$$

where $\tilde{r}(\nu) = \frac{\tilde{n}(\nu)-n_{air}}{\tilde{n}(\nu)+n_{air}}$, $\tilde{t}_{12}(\nu) = \frac{2n_{air}}{\tilde{n}(\nu)+n_{air}}$ and $\tilde{t}_{21}(\nu) = \frac{2\tilde{n}(\nu)}{\tilde{n}(\nu)+n_{air}}$ are the complex Fresnel coefficients at normal incidence; m is the order of reflection in the time-domain (TD) response and can be derived through: $t_{max} \geq \frac{n}{c}d(1+2m)$, with t_{max} being the signal duration in TD (normally of the order of 10s of ps).

Due to the nature of the TF (eqn. 3.5), a non-linear regression algorithm must be adopted to estimate $\tilde{n}$. The recursive fitting procedure involved requires initial (or 'seed') guesses to be made for n and k. Regression fitting iteratively refines until the experimental TF approaches to within a pre-defined threshold of identity with the model or theoretical TF.

3.3 Dependence of estimates of n and k on initial guesses

The difference between theoretical and experimental TFs will serve as an objective function (OF):

[2] A short scan is considered to be shorter than the time of flight for the first internal reflection.

$$\widetilde{OF}(\nu) = \tilde{H}_{th}(\nu) - \tilde{H}_{exp}(\nu) \tag{3.6}$$

Roots of this function will represent the actual optical constants of the sample. Considering that both TFs are complex, the equality-condition of two complex numbers results in two equivalent pairs of equations. One concerns the equality of real and imaginary parts of both TFs and the second for the modulus and argument. Let us consider first the simpler case of equating real and imaginary parts of the TFs. This approach shows that the roots of the OF exhibit oscillatory and poly-root behaviour when calculating optical constants. Consequently regression analysis may converge to different roots depending on the initial guesses made. The periodicity of the possible roots of the OF is governed by the following relation (derived from complex TF of eqn. (3.5)):

$$T_{osc} = \frac{c}{\nu d}, \tag{3.7}$$

d again is the depth of the sample.

The period T_{osc} is dimensionless; it is used to describe the function of dimensionless argument, i.e. refractive index. In order for the iteration procedure to converge to the right root, the starting point has to be specified within the range

$$n_{actual} \pm T_{osc}/2. \tag{3.8}$$

In most cases initial-guess values for n and k are respectively derived from the delay in the sample-response and the attenuation in response amplitude [Pupeza'07, Krüger'11]; thus

$$n_{est} = n_{air} + c\Delta t/d \tag{3.9}$$

and

$$k_{est} = -\frac{c}{2\pi\nu d} ln\left(\frac{E_{s,max}}{E_{r,max}}\frac{(1+n)^2}{4n}\right). \tag{3.10}$$

Δt is the delay of the main–pulse transmitted through the sample with respect to reference pulse; $E_{s,max}$ and $E_{r,max}$ are the maximum amplitudes of sample and reference spectra.

Estimation of n and k based on a discrete initial guess for each frequency point is problematic when investigating dispersive media. An example demonstrating the failure of an advanced algorithm to correctly estimate optical constants is illustrated using a 3 mm thick sample of z-cut quartz. The refractive index of this material is slightly

dispersive in the THz frequency domain. It ranges from 2.11 to 2.15 between 0.1 and 3.8 THz (Fig. 3-1). The curve obtained from the advanced method is smoother and consequently more physically significant, but regression analysis converges to a false root of the OF at around 2.8 THz. This cut-off frequency can be validated from eqn. (3.7) and equal to $0.5T_{osc}$, which in this case is 0.018. This corresponds to the difference between the initial guess and the actual value of n at the cut-off frequency. These artefacts can be mistakenly attributed to the dynamic range limitation of the system.

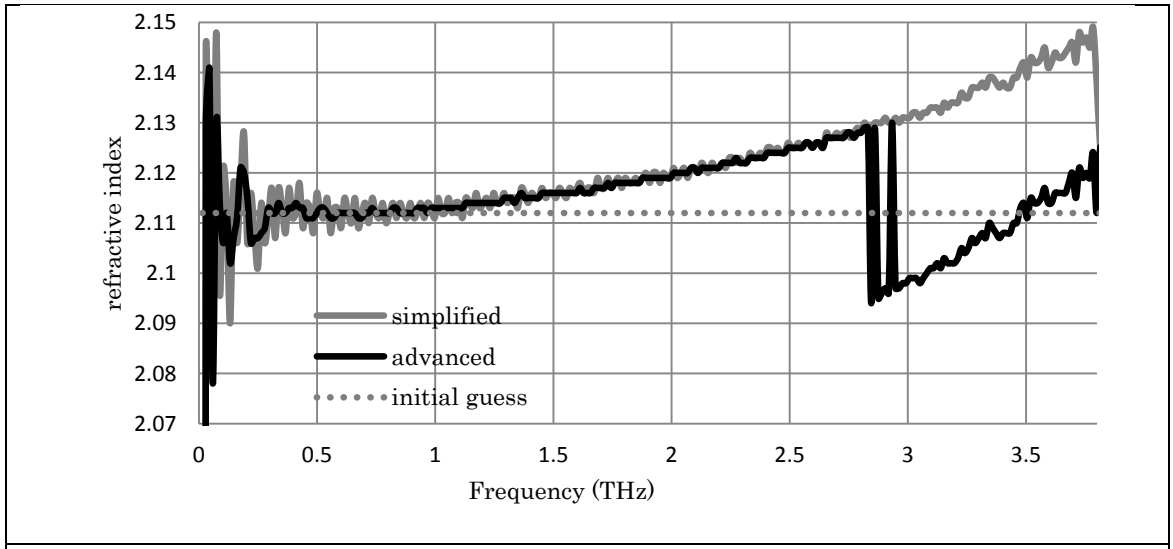


Figure 3-1. Refractive index calculation for a 3 mm thick z-cut quartz sample. The graph shows estimation of n using the simple transfer function (eqn. 3.2) and the more rigorous transfer function (eqn. 3.5). The initial guess for estimating n by eqn. (3.5) is shown as a straight line.

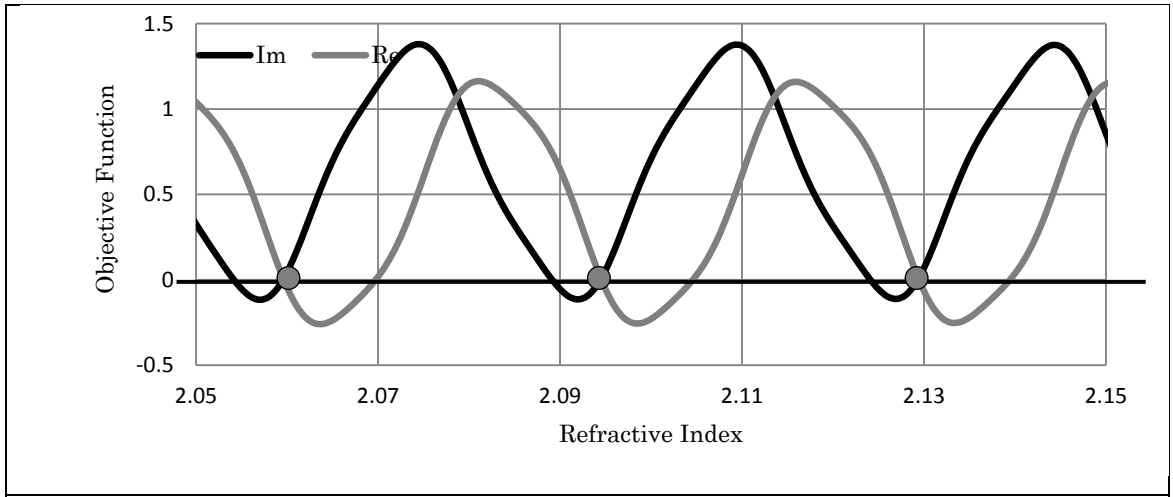


Figure 3-2. Possible values of n are shown (gray circles) for a 3 mm thick sample of z-cut quartz when real and imaginary parts of $H_{th}(\nu)$ and $H_{exp}(\nu)$ are respectively equated.

The dependence of $OF(\nu)$ roots on n and k can be analysed separately. For the 3 mm thick z-cut quartz, Fig. 3-2 shows the possible roots when real and imaginary parts of $H_{th}(\nu)$ and $H_{exp}(\nu)$ are respectively equated.

However, when the modulus and argument of both TFs are respectively equated, the poly-root problem is avoided. The modulus of OF still provides multiple roots but the unwrapped phase of the OF limits it to a single possible root (Fig. 3-3). This condition forces the iterative procedure to converge to the correct value of refractive index.

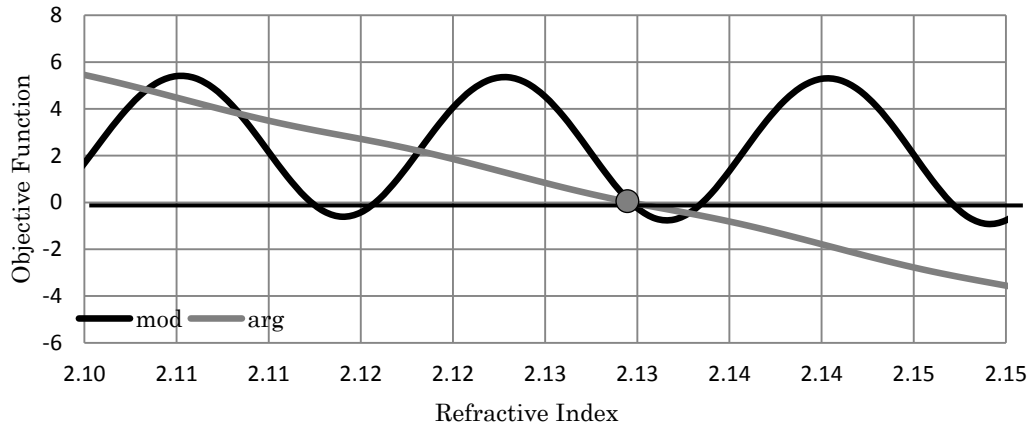


Figure 3-3. For a 3 mm thick z-cut sample of quartz: dependence of the objective function modulus (oscillatory trace) (scaled by 50 times for clarity), and unwrapped phase (straight line) of refractive index. Both traces were generated at 1 THz.

The validity of the procedure for estimating n, based on the modulus and unwrapped phase of the transfer functions, is demonstrated on the highly dispersive material – lithium niobate ($LiNbO_3$). The estimation of optical constants by the three different approaches is shown in Fig.3-4. Values obtained by using real and imaginary parts of the TF (eqn. 3.5) have a cut-off at around 1.3 THz, while the modulus and argument approach predicts valid results over the considered frequency domain. Accounting for internal reflections yields smoother dispersion curves in plots of n and k. Fig. 3-4B demonstrates that at lower frequencies the estimates of k can be several times different when using simple and enhanced TF.

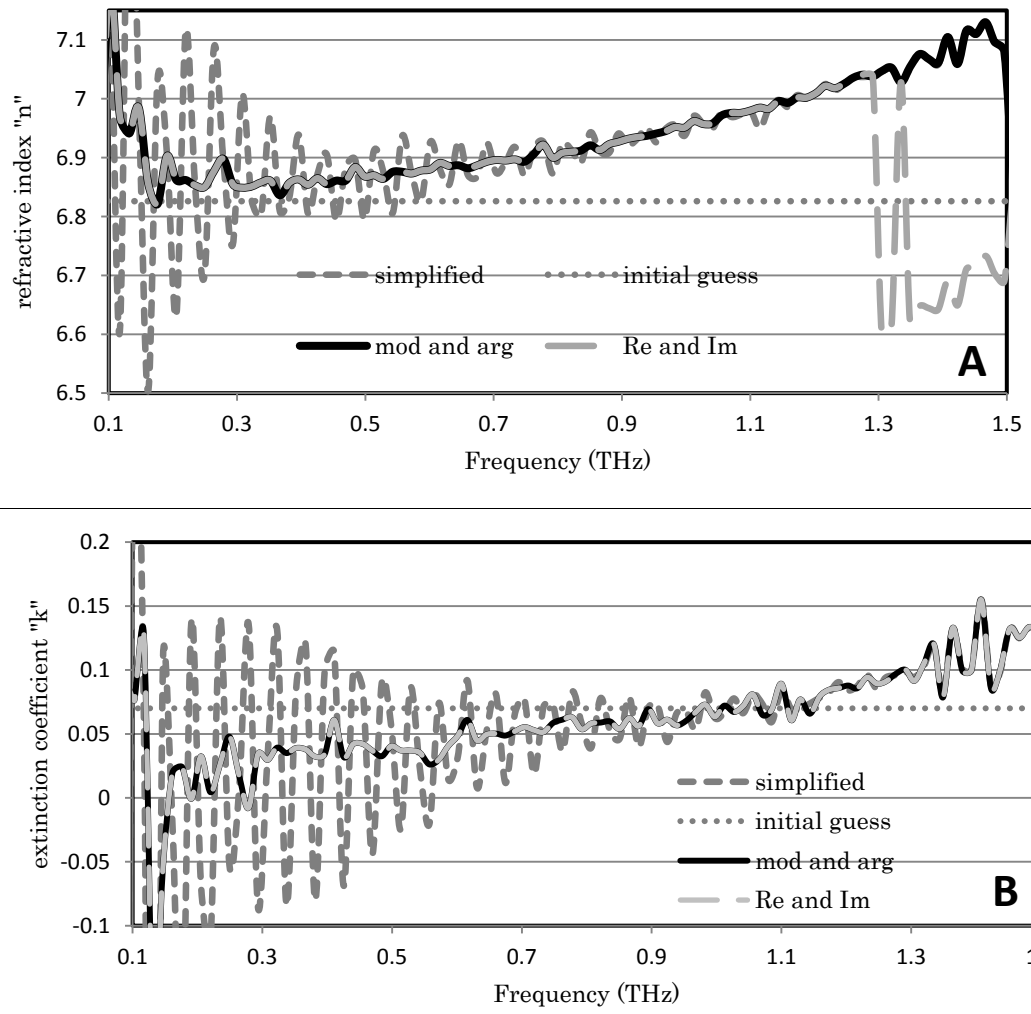


Figure 3-4. A: shows estimates of$n$ for $LiNbO_3$ using: 1) the simplistic approach (eqn. 3.2); 2) the more rigorous approach of eqn. (3.5) that accounts for multiple internal reflections, the black trace by equating modulus and argument between theory and experiment and the light grey, by equating real and imaginary parts; 3) The grey dotted trace shows the initial guess used during the rigorous approach. A $LiNbO_3$ sample of 0.5 mm thickness was used for these measurements. B shows the same but for k.

The differing uses of the same OF yield different results. Exploiting real and imaginary parts of TFs fails at high frequency even for weakly dispersive materials like z-cut quartz. But by using modulus and argument instead, the global minimum is found (i.e. correct root of the OF), even for high-dispersive materials like $LiNbO_3$. Again the phase of the TF have to be unwrapped in order to result in a single-root behavior of the OF (Fig. 3-3).

Similarly the dependence of the OF roots on the initial guess can be overcome by introducing a step-dependent initial guess. Meaning that initial guess for each frequency

point is the actual parameter value obtained for previous step. The algorithm has to start from the lowest operating frequency where the period of OF is the highest. However this approach requires additional programming effort and is not described in detail here.

3.4 Fabry-Perot Effect

The main difference between the simplistic and advanced methods of estimating $\tilde{n}$ is based on how multiple internal Fabry-Perot (FP)-like reflections are formulated. Approaches for interpreting this effect vary. The approach of Naftaly *et al.* was simply to removing spurious oscillations in the THz spectra [Naftaly'07b]. It does so by treating the FP oscillations as a primary peak convolved with delta functions. The constraints of this algorithm are that samples must have low absorption and dispersion and must be sufficiently thick to avoid overlapping of main and secondary peaks. Duvillaret *et al.* otherwise seek to eliminate the FP or etalon effect [Duvillaret'96]. It starts from approximate values for n and k to estimate the FP optical 'ringing'. The complex TF is then divided by this term and after several iterations the ringing is significantly reduced. The drawback of this method is that it requires a formulation to approximate the ringing. This approximation introduces additional error to the final material parameters and will be described later.

A general analysis [Born'99], allows for multiple reflections within a material slab enveloped in a medium of differing wave impedance; thus,

$$\begin{aligned}\widetilde{FP}(\nu) &= \left\{1 + \tilde{r}^2(\nu)\tilde{A}^2(\nu) + \tilde{r}^4(\nu)\tilde{A}^4(\nu) + \cdots\right\} = \sum_{i=0}^{\infty}\left[\tilde{r}^2(\nu)\tilde{A}^2(\nu)\right]^i \\ &= \left[1 - \tilde{r}^2(\nu)\tilde{A}^2(\nu)\right]^{-1}.\end{aligned} \tag{3.11}$$

Eqn. (3.11) represents the approximation for the infinite number of multiple internal reflections. The TF (eqn. 3.5), incorporating eqn. (3.11) now becomes:

$$\widetilde{H}_{Approx.}(\nu) = \frac{\tilde{t}_{12}(\nu)\tilde{t}_{21}(\nu)\exp\left(-\frac{i2\pi\nu d(\tilde{n}(\nu) - n_{air})}{c}\right)}{1 - \tilde{r}^2(\nu)exp\left(-\frac{i4d\pi\nu\tilde{n}(\nu)}{c}\right)}. \tag{3.12}$$

A lot of research groups [Kruger'11, Duvillaret'96, Duvillaret'99, Withayachumnankul'05] use this type of TF in estimating $\tilde{n}$. Although eqn. (3.12) is readily used, it introduces an additional error to estimates of $\tilde{n}$ due to the assumption of there being an infinite number of internal reflections.

In practice, however, the ringing is limited due to the finite energy of the primary incident electric field. As such, eqn. (3.11) takes the form of:

$$FP_0(\nu) = \sum_{i=0}^{m} [r^2(\nu)A^2(\nu)]^i, \tag{3.13}$$

where m represents the order of internal reflection.

The number of copies of the primary incident pulse due to ringing is 1 to 3 in the majority of THz-TDS measurements. This varies with the optical thickness and absorption of a sample. In this case analysis using eqn. (3.12) introduces noticeable errors. For accuracy, then, the order of internal reflection should be determined (even if not evident in the TD response), accompanied by application of the TF of eqn. (3.5).

We now quantify these errors and account for them in the total uncertainty associated with an estimate of $\tilde{n}$. A 1 mm thick silicon plate is employed for this. Reference and sample responses are shown in Fig. 3-5. The figure circles two near-copies of the primary (incident) pulse.

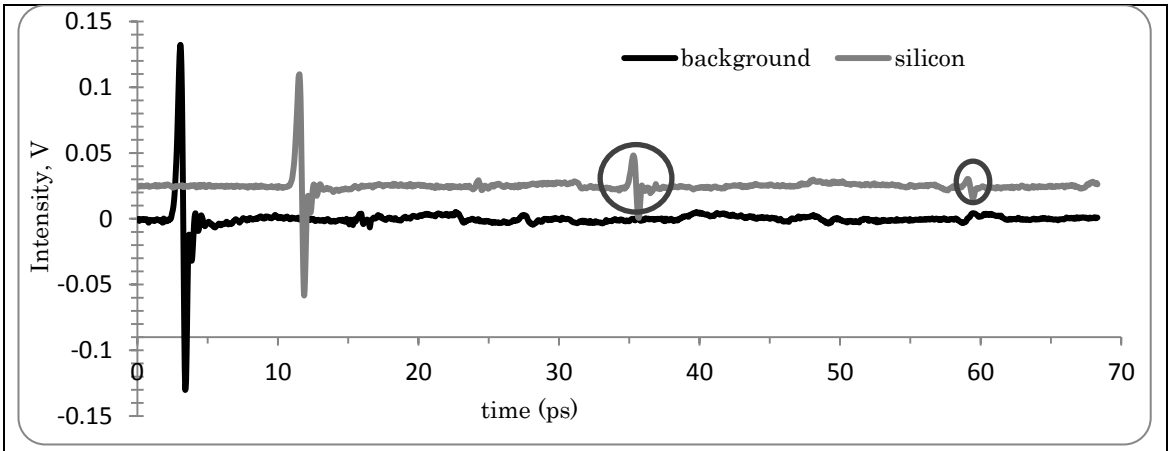


Figure 3-5. Time domain response of the reference path (black trace) and the path when filled with a 1mm thick plate of silicon plate (grey trace). Silicon values are offset by +0.025 for clarity. Measurements were done under normal conditions of atmosphere and temperature. Two internal reflections are circled.

Fig. 3-6 shows the uncertainties introduced by using the approximated TF (eqn. 3.12). The dark grey curve corresponds to additional uncertainty when no ringing is present in the time-domain response of the sample. This was achieved by taking a scan-time of 35 ps. It then shows the additional uncertainty when one reflection is present in the time-

domain response (scan-time 58 ps). The light grey curve accounts for two internal reflections (i.e. the whole scan). The approximate treatment of eqn. (3.12) can be seen to be best suited to measurements with multiple reflections (i.e. >2), this holds for optically thin samples.

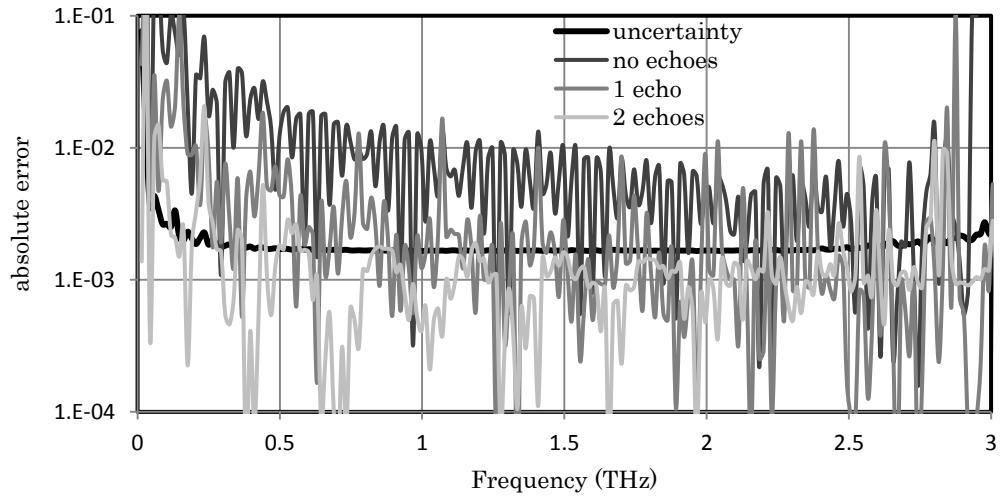


Figure 3-6. The additional errors in estimation of n resulting from the analysis procedure of eqn. (3.12) when the sample response contains: 1) no ringing (dark grey); 2) one reflection (grey); 3) two internal reflections (light grey). The black solid-curve represents uncertainty in the refractive index when eqn. 3.5 is used for analysis.

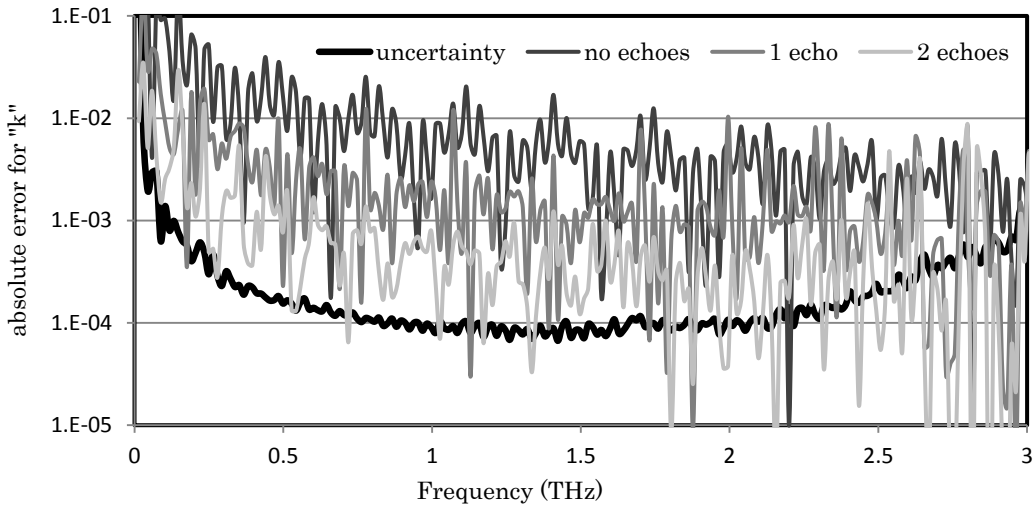


Figure 3-7. The additional errors in estimation of k resulting from the analysis procedure of eqn. (3.12) when the sample response contains: 1) no ringing (dark grey curve); 2) one reflection (grey); 3) two internal reflections (light grey). The black solid-curve represents uncertainty in the refractive index when eqn. 3.5 is used for analysis.

For cases where the number of internal reflections in a time-domain response are no more than 2 this approach is no longer reliable, since the uncertainty increases at some frequencies by one-to-two orders of magnitude, as compared to the exact procedure of eqn. (3.5), where a finite number of reflections is considered (black curve of Fig. 3-6). Fig. 3-7 further shows additional uncertainties as in Fig.3-6, but for k instead. Notice the additional uncertainty, due to the FP approximation of the TF of eqn. (3.12), is relatively greater here than for n. This additional term needs to be introduced to uncertainty calculations, and can be evaluated as follows [Withayachumnankul'08a]:

$$f_n(\nu) = \frac{c}{2\pi\nu d}\left|f_{arg}(\nu)\right| \tag{3.14}$$

and

$$f_k(\nu) = \frac{c}{2\pi\nu d}\left[\left|f_{mod}(\nu)\right| + \frac{1}{n(\nu)}\frac{n(\nu)-n_{air}}{n(\nu)+n_{air}}\left|f_n(\nu)\right|\right], \tag{3.15}$$

where

$$f_{arg}(\nu) = arg\left(H_{approx}(\nu)\right) - arg\left(H_{exact}(\nu)\right) \tag{3.16}$$

and

$$f_{mod}(\nu) = ln\left|H_{approx}(\nu)\right| - ln\left|H_{exact}(\nu)\right| \tag{3.17}$$

3.5 Error analysis for the extraction procedure

An uncertainty analysis has utility in identifying and controlling sources of systematic and random error in estimating complex refractive index. Gaussian error propagation models have been used here. The accuracy of the optical constants is an essential supplement to material parameter extraction. It states how accurate the final results are and usually helps to identify ways of minimizing the errors from different sources. Starting from variance in the real and imaginary part of the sample and reference spectra the uncertainties of material parameters are derived [Pupeza'07, Krüger'11]. A substantial analysis of uncertainties introduced by random errors was done by Duvillaret *et al.* [Duvillaret'00]. The most comprehensive and well-grounded study of material parameters' uncertainty was made by Withayachumnankul *et al.* [Withayachumnankul'05,08a]. They quantify a wide range of sources of error: random and systematic variance of signal amplitude; influence of errors in sample thickness; approximations made in constructing TFs; sample alignment and even systematic errors in physical constants such as the refractive index of air. Following such a procedure, my aim is to eliminate or reduce the influence of as many sources of error as identifiable in

order to achieve highest accuracy, and this will also include, and appropriately account for, signal ringing within the sample.

The basic approach to evaluating the measurement error is just based on the estimation of standard deviation of the final quantity. However, this method ignores the systematic errors that are normally present in any measurement, and cannot distinguish between different sources of uncertainties.

There is significant variation across groups in treating errors. The majority of THz groups [Pupeza'07, Krüger'11] use Gaussian error propagation modelling in order to estimate the uncertainty of material parameters. They employ:

$$u_{n,k}(v) = \sqrt{s^2_{n,k\_samp} + s^2_{n,k\_ref}} \tag{3.18}$$

where $u_{n,k}(v)$ is the uncertainty of n or k; s^2_{n,k_samp}, s^2_{n,k_ref} are respectively variances in n and k due to variances in the sample and reference measurements. Knowing the variances of signals in TD assumes that repeated measurements of sample are available and the averaged trace is used for extraction. Withayachumnankul *et al.* [Withayachumnankul'08a] however follows the GUM[3] guidelines on error propagation modelling by further dividing the variances in sample and reference responses with the number of measurements performed. This leads to a smaller uncertainty interval and allows localizing the mathematical expectation of n and k if the number of measurements tends to infinity. According to this model, the final uncertainty is formulated as:

$$u_{n,k}(v) = \sqrt{\frac{s^2_{n,k\_samp}}{N_{samp}} + \frac{s^2_{n,k\_ref}}{N_{ref}}} \tag{3.19}$$

where N is number of measurements performed.

This discrepancy in how errors are treated results in misinterpretation on levels of uncertainty of n and k. For instance, the difference in uncertainty levels resulting from time-domain amplitude variances can be one order of magnitude if a hundred measurements are considered.

[3]Recommendation by the International Organisation of Standards (ISO) known as *"Guide to the Expression of Uncertainty in Measurements"* (GUM).

Following the guidelines of Withayachumnankul *et al.* [Withayachumnankul'05,08a] seven sources of errors are analysed and listed below where the last one is the unique addition of this chapter.

1. Random and systematic variance of E_{Field} amplitude (incorporating emitter noise, detector noise and random fluctuations in laser amplitude).
2. Sample thickness

2a. Error due to resolution limit of thickness measuring device (micrometre).

3. Approximations of the TF due to using only the real part of the refractive index in Fresnel formulae.
4. Sample alignment[4] (sample tilt leading to off-normal incidence of the THz beam).
5. Systematic errors in the refractive index of air.
6. Neglecting appropriate treatment of multiple internal reflections within the sample (using eqn. (3.3, 3.4)).
7. Using eqn. (3.12) instead of eqn. (3.5) when dealing with high-order ringing.

Parts 6 and 7 in the list contribute to the final uncertainty only when using the simplified extraction procedure. For optical constant estimation using the TF of eqn. (3.12) – part 6 can be omitted. Terms 6 and 7 can be calculated using eqn. (3.13), (3.14), other terms is calculated according to [Withayachumnankul'08a].

The parts in a Gaussian error propagation model for uncertainty, sum as follows:

$$u_{n,k}(v) = \sqrt{\underbrace{\frac{s_{n,k}^2}{N_{meas}}}_{1} + \left(\underbrace{\frac{s_{n,k_d}^2}{N_d}}_{2} + \underbrace{s_{n,k_\delta}^2}_{2a}\right)} + \underbrace{f_{n,k_{TFA}}}_{3} + \underbrace{f_{n,k\_align}}_{4} + \underbrace{f_{n,k\_n0}}_{5} + \underbrace{f_{n,k\_FP}}_{6} + \underbrace{f_{n,k\_FPE}}_{7} \quad (3.20)$$

The use of these separate parts, however, depends on measurement procedure and parameter estimation method. When the general simplistic solution of eqns. (3.3) and (3.4) are used, all terms have to be considered in the calculation of the final uncertainty. This would compute the highest level of uncertainty in n and k, as shown in Fig. 3.8. When an enhanced extraction procedure is performed using eqn. (3.5), combined with numerical thickness determination, then only parts 1, 2 and 4 should be considered for evaluating the total uncertainty. In the case of numerical thickness determination, as in the algorithm employed by Pupeza *et al.* [Pupeza'07], the 2a uncertainty term is

[4] Alignment is calculated with a $\pm 2^{\circ}$ accuracy of normal incidence of THz beam on the sample.

omitted. Finally, when eqn. (3.12) is used, parts 1, 2, 4 and 7 need to be considered in estimating uncertainty since part 7 represents the additional error from approximation of ringing. This part however becomes negligible when more than five internal reflections are present in the time domain response of the sample under test. Table 3-1 summarizes the usage of error terms for different extraction procedures.

Table 3-1. The use of error terms

	Using TF eqn. 3.2	Using TF eqn. 3.12	Using TF eqn. 3.5
Error terms to be considered	All	1, 2, 4, 7	1, 2, 4

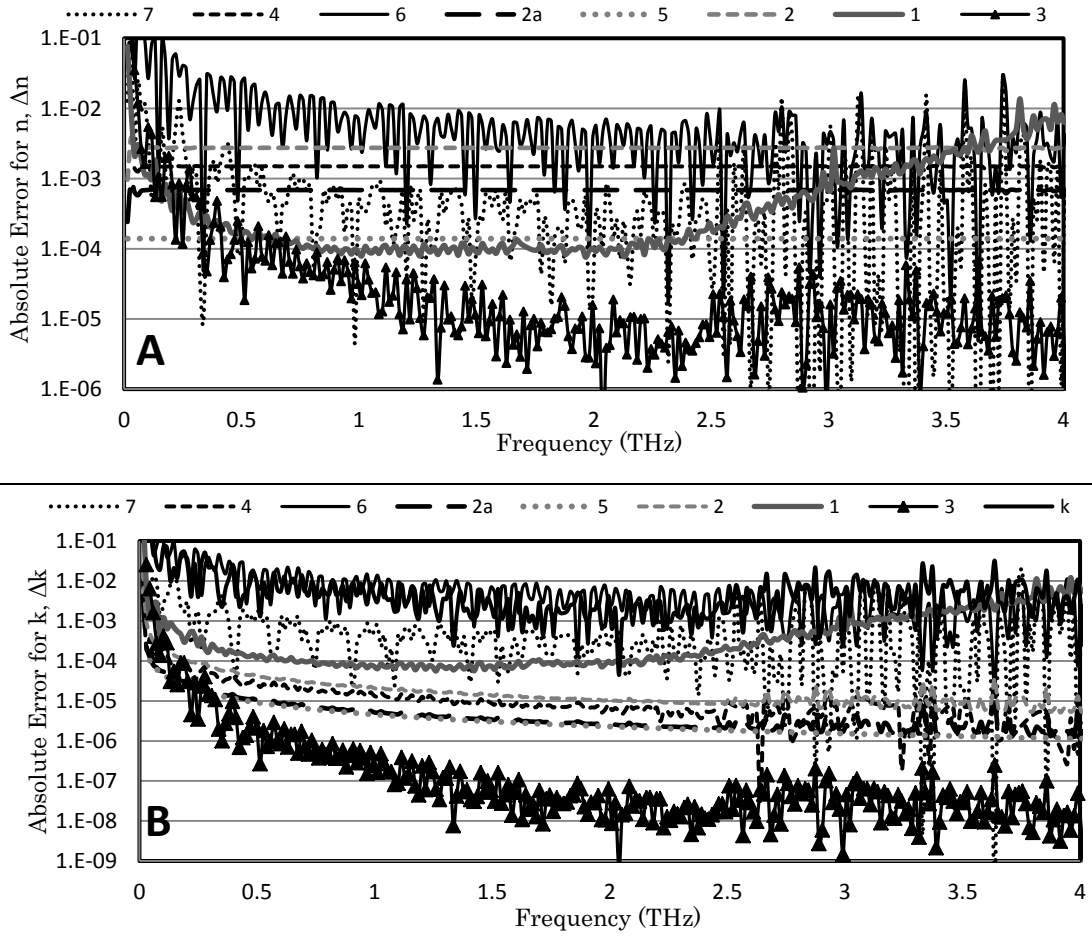


Figure 3-8. A: shows the contribution of each error part to the total uncertainty in n for the simplified extraction procedure. B shows the same for k.

Fig. 3-8 shows the contribution of each constituent error term to the total uncertainty in n and k of silicon when eqns. (3.3) and (3.4) are used as an analysis procedure.

The level of random and systematic error in E_{Field} (1st error term for n and k) is almost the same for both *n* and *k*, but k itself is usually several orders of magnitude less than n so, relatively, it is more sensitive to s_k^2. Fig. 3-8 further shows that ringing (6th and 7th error terms) has about equal contribution towards the total uncertainty in n and k. Note that the error caused by the ignoring of ringing (f_{n,k_FP}) is even higher than the estimate of k itself, throughout the considered frequency domain.

From this one can conclude that the dominant factor in the uncertainty of n and k are the $s_{n,k}^2$ variances and the ringing (f_{n,k_FP} and f_{n,k_FPE}), while n is also heavily influenced by sample thickness and sample alignment errors.

The contribution of other error terms considered in Fig. 3-8 to the total uncertainty in n and k is lower by more than two orders of magnitude compared to their respective optical constants. The least contribution to the total uncertainty is presented by the 3rd term (assuming purely real Fresnel coefficients instead of complex) due to low imaginary part of refractive index. For strongly absorbing materials this term becomes more dominant.

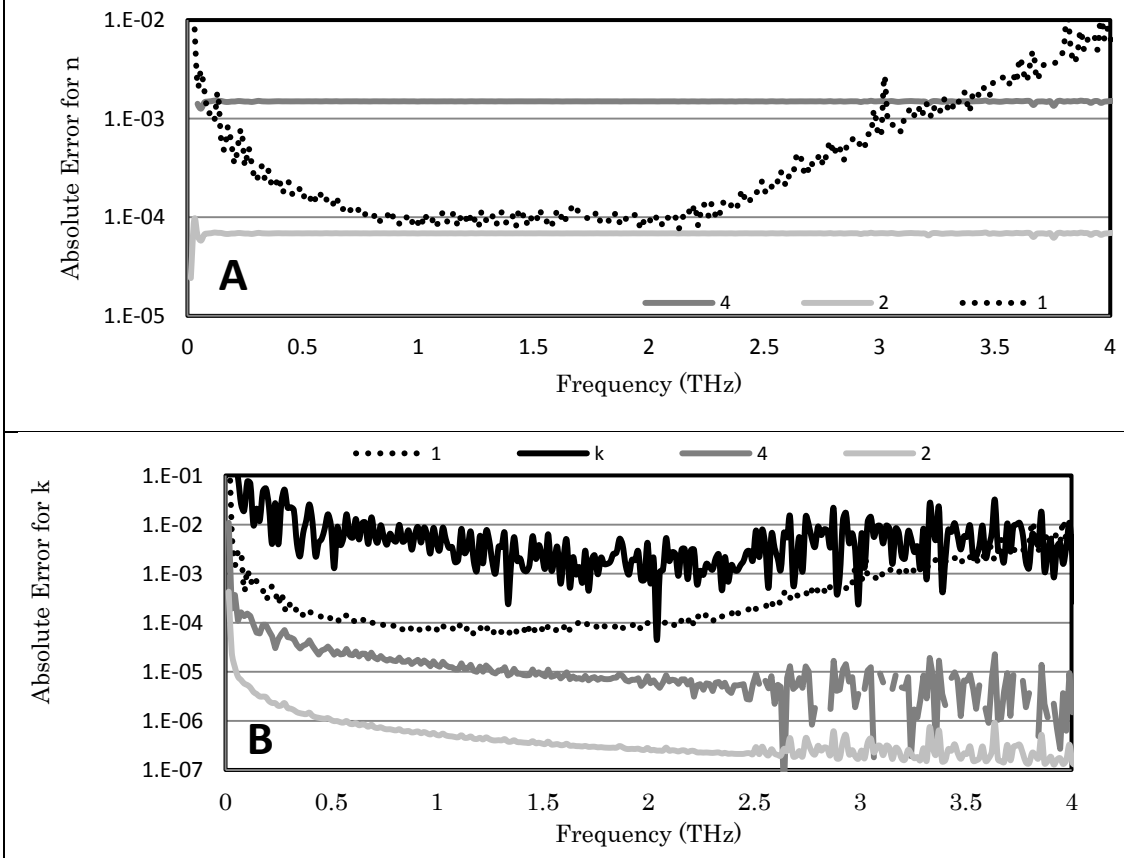


Figure 3-9. A: shows the contribution of each error term to the total uncertainty in n for the advanced extraction procedure. B shows the same for k.

Fig. 3-9 shows the error terms contributing to uncertainty when eqn. (3.5) is used for evaluating n and k. Recall this enhanced extraction procedure excludes some error terms (refer to Table 1). Only $s_{n,k}^2$ variances, thickness and sample alignment errors, need be considered. Since the f_{n,k_FP} and f_{n,k_FPE} terms no longer affect the accuracy of the optical constants, the E_{Field} variance (s_k^2), is shown to have the most significant effect on k. The values of errors resulting from s_k^2 are about two orders of magnitude less than those of k through most of the considered frequency domain. For n the alignment of the sample (f_{n,k_align}) has the highest impact on the final uncertainty (Fig. 3-9A). Alignment is calculated as in [Withayachumnankul'08a] with the assumption of a $\pm 2^{o}$ accuracy of normal incidence of THz beam on the sample. Fig. 3-9 also provides insights into the frequency distribution of the signal noise (trace 1). The noise curve is in fact proportional to the inverse of the sample spectrum, i. e. 1–2 THz domain corresponds to the minimum of the noise and at the same time to the maximum of spectral amplitude; above 2 THz the spectrum starts to decay while the noise increases. On the other hand in the time-domain the distribution of noise is less uniform. The highest signal fluctuations are concentrated at the position of the primary THz pulse (it can be even 3-4 orders higher compared to the rest of the trace), which can be seen by calculating the standard deviation of the time-domain response of the sample.

Another source of error in the material parameters is pseudocoherence – the interferometric variation in received amplitude as a function of surface roughness and frequency [Hadjiloucas'99]. Different beams propagating through the sample can travel slightly different path-lengths that can lead to a destructive interference and reduction in the observed signal. The contribution of pseudocoherence is proportional to frequency and inversely proportional to sample thickness. This error is considered to be negligibly small in our study since the surface roughness of most samples investigated is much less than the operational wavelength.

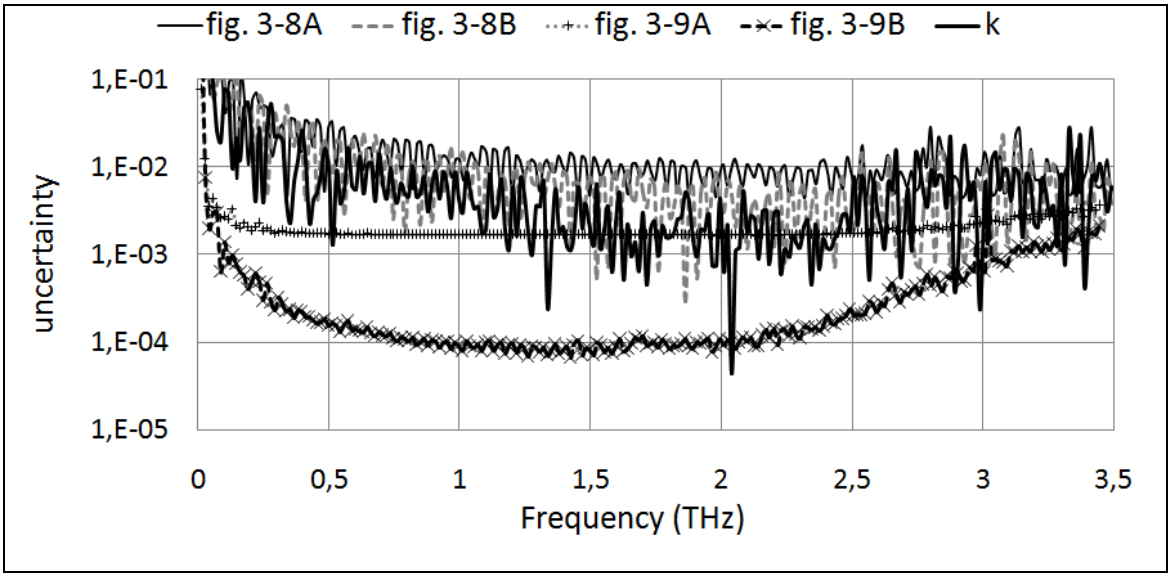


Figure 3-10. This graph shows the total uncertainty in n for general (fig. 3-8A) and enhanced (fig. 3-9A) analyses. The same applies for k (fig. 3-8B and fig. 3-9B respectively). The black heavy curve shows values of k for comparison purposes.

Fig. 3-10 shows the difference in total uncertainty for the simplistic evaluation procedure of complex refractive index eqns. (3.3) and (3.4), and the advanced of eqn. (3.5). Uncertainties in n following eqn. (3.3) and (3.4) are generally one-to-two orders of magnitude higher than for eqn. (3.5). For n the absolute difference in uncertainty between two methods is quite high (one order of magnitude), but relatively the improvement is not so significant – from 1% to 0.06%. But for k the difference between the two approaches is striking. A simplistic analysis yields a total uncertainty Δk on the order of k itself. Advanced analysis reduces it to 3-5 % of k. This holds for silicon plate studied here and other materials that have well-pronounced FP-echoes in TD response. This is a significant improvement in accuracy, as absorption properties are of particular interest in bio-molecular studies. The extinction coefficient, therefore, is a parameter limiting the viability of the simpler approach. There is little additional computational expense in employing the advanced analysis and it is ultimately preferred for correctly accounting for multiple internal reflections.

3.6 Potential issues that influence optical constants extraction

3.6.1 Sample preparation

Sample preparation is a critical step in measurement repeatability. THz-TDS is sensitive to structural changes of material or to fillers/impurities introduced in preparation. Choice and amount of fillers are critical to optimum measurement. In terms of reducing

scattering effects it is better to use a filler material whose refractive index is as close as possible to that of the sample and particle size as small as possible [Federici'05].

The main materials which are often used as fillers for THz spectroscopy are polyethylene (PE) and Teflon (PTFE). Both absorption and index of refraction will differ from the intrinsic sample parameters; which should be properly taken into account during data analysis [Franz'08].

For accurate parameter extraction the purity and humidity of both sample and filler have to be considered. Sample thickness too is significant. A thick sample (thickness is greater than 10 times the wavelength at 1 THz) will cause higher signal attenuation and will result in a limited frequency range for absorption coefficient determination. On the other hand, a thin sample behaves as a Fabry-Perot cell [Pupeza'07].

3.6.2 Alignment

Fine alignment of THz-TDS optical and THz paths is crucial for quantitative measurements. Alignment is most critical for the first optical components in the beam-path. A small error in beam-angle at the first mirror can lead to several millimeters down-beam misalignment. For example, rotating the first plane mirror in our THz-TDS system around its axis by only 0.5 mrad (which corresponds to a 5 μm shift of the mirror edge) can cause 1 mm beam displacement at the photoconductive antenna. This can cause significant signal drop and decrease in operational bandwidth, up to complete signal losing. Fig. 3-11 illustrates the measured signal drop on the detector output depending on the displacement of the GaAs antenna off focus from the first off-axis parabolic mirror collecting THz radiation.

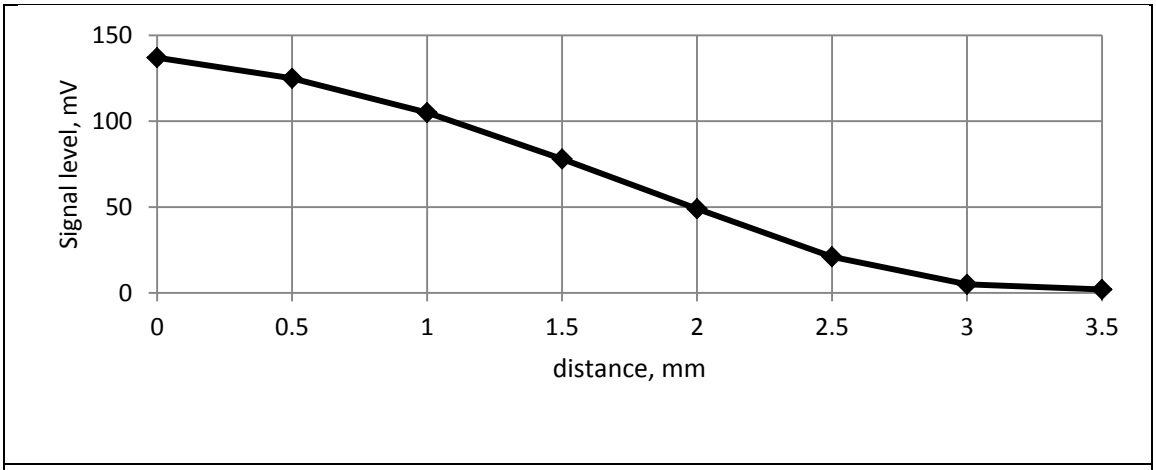


Figure 3-11. Dependence of signal level on photoconductive antenna displacement.

3.6.3 DR limitations and resolution of the system

This section intends to illustrate the potential limitations in absorption coefficient measurements. Neglecting the Fabry-Perot effect, the simplest formula for the sample absorption coefficient can be derived from Beer-Lambert law [Jepsen'05]:

$$\alpha = -\frac{2}{d}\ln\left[\frac{E_{sample}}{E_{reference}}\frac{(n+1)^2}{4n}\right] \tag{3.21}$$

where E_{sample} and $E_{reference}$ are spectral amplitudes for sample and background respectively. Inspection of (3.21) shows that the absorption coefficient mainly depends on the relation $E_{sample}/E_{reference}$ which is proportional to dynamic range (DR). Thickness d is constant and refractive index changes are not strong enough to make a significant contribution to absorption coefficient (3.21) for most materials in the THz frequency range. As a result, the operating frequency domain for THz-TDS normally extends up to 4 THz, but for measuring highly absorbing or thick samples the upper limit for absorption occurs below this. This is demonstrated by the measurement of two glycine pellets of different thickness on our TDS. As expected, significant values of absorption for thinner sample can be measured over a broader frequency domain– up to 1.25 THz, while for 2.5 mm thick glycine pellet – only up to 0.9 THz (Fig. 3-12). The same effect was observed by Jepsen et al. [Jepsen'05]. So, for every type of material under test there exists a maximum measurable absorption coefficient which directly depends on the DR of the system.

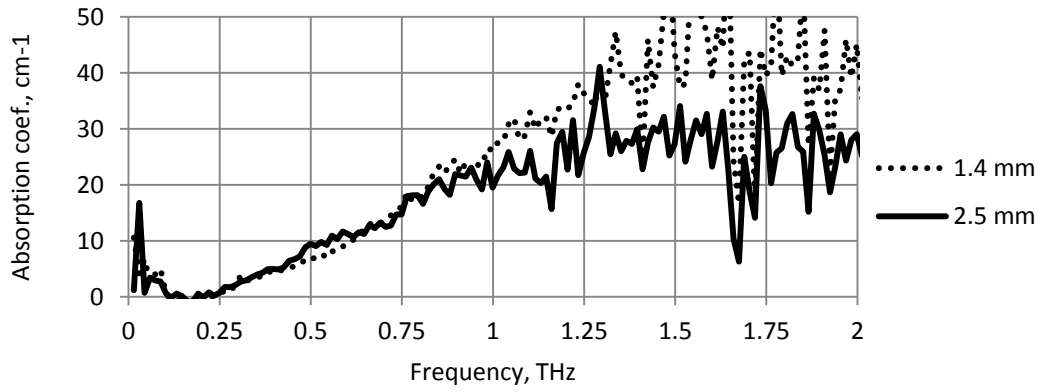


Figure 3-12. Absorption coefficient of glycine pellets of 1.4 mm and 2.5 mm thicknesses

The complex refractive index spectrum of a material is obtained by FFT of the recorded temporal waveform. From Fourier theory, the resolution in the frequency domain is inversely proportional to the signal duration in the time domain. So the direct

way to enhance the frequency resolution is to increase the scanning length of the delay stage. The penalty, however, is reduction in dynamic range, simply because a longer time domain accrues more noise. There is a tradeoff between resolution and dynamic range of the system [Mickan'04].

3.6.4 Humidity

An important factor to be considered in THz-TDS measurements is water vapor absorption within the sample chamber. Among other gases, water vapor exhibits a very strong rotational absorption spectrum in the terahertz range [Zeitler'07]. Background measurements performed in our THz laboratory under NTP (50% humidity) and with the THz box purged with nitrogen (to bring about a 3% humidity) are shown in Fig. 3-13. The water vapor absorption peaks can be easily seen, as well as their decreasing for lower humidity. It is important to know the resonance frequency of these peaks in order to avoid confusion with sample absorption features. These sharp absorption peaks can also lead to discontinuities in the phase of the sample, further resulting in sharp features in rather flat curves of material optical properties. To minimize the contribution of water vapor to the sample spectrum, the sample chamber is either purged with dry air, nitrogen or evacuated throughout the measurements. If these options are not available, the algorithm for elimination of water vapour absorption was developed by Withayachumnankul *et. al.* [Withayachumnankul'08b]. They study a computational means for addressing the problem arising from water vapour absorption. Initially, the complex frequency response of water vapour is modelled from a spectroscopic catalogue. Using a deconvolution technique, together with fine tuning of the strength of each resonance, parts of the water vapour response are removed from a measured THz beam signal, with minimal signal distortion, thus providing an alternative for purging the measurement box.

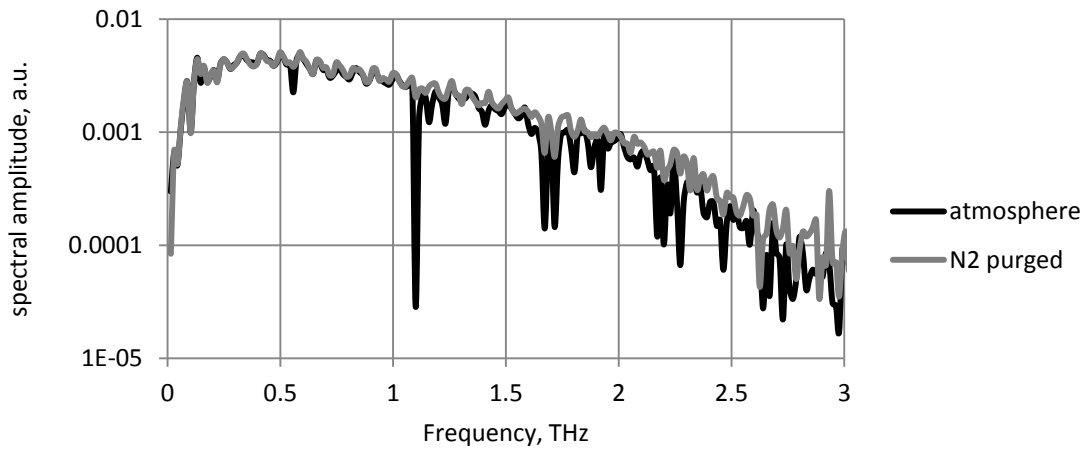


Figure 3-13. Absorption spectrum for air at different humidity

3.7 Numerical thickness determination algorithm

Thickness of the sample is known to be the major source of error in THz-TDS [Duvillaret'99], however accurate mechanical thickness determination is not always possible. Therefore the first step to maximize the accuracy of the refractive index is to determine the effective thickness of the material under test in a precise manner.

Various methods have been employed for optical (read also, 'numerical'), thickness determination. Kruger *et al.* [Krüger'11] used two measurements of the same material but having different thicknesses. After extracting all possible solutions via TF regression analysis, coincidence of the two curves corresponded to the actual thickness. The disadvantage of this method is a necessity to reproduce identical conditions for both measurements. It also has a relatively high computational cost. The approach has been developed for determination of thickness for sub-100-μm samples [Scheller'09a]. Its algorithm relies on an additional Fourier transform of the frequency-dependent material parameters to a quasi-space domain. Other methods [Scheller'09b] are based on time-domain reconstruction of the THz response, but they too suffer from high computational requirements. Here a TV method is adopted as a simple and reliable approach for numerical thickness determination [Dorney'01]. This method employs the complex nature of an advanced TF (eqn. 3.5), that is fortuitous in assisting estimation of optical thickness in a very accurate manner. The principle of TV is based on the fact that in most cases the dispersive curves of material properties are smooth, so that an incorrect determination of thickness manifests itself as spurious oscillations (Fig. 3-14).

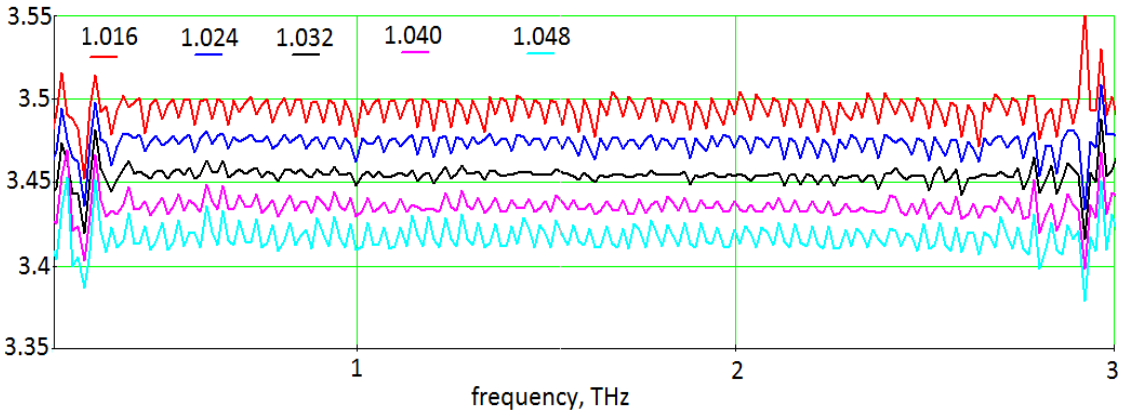


Figure 3-14. Refractive index of Silicon plate, plotted for different thickness values in mm.

By iteratively varying thickness, the algorithm selects the best-behaved curve, and this corresponds to the most accurate thickness estimate. The value of TV is calculated for each thickness as [Pupeza'07]:

$$TV = \sum_{\nu_m}^{\nu_l} (|n_k - n_{k-1}| + |k_k - k_{k-1}|), \tag{3.22}$$

where ν_m to ν_l defines the frequency domain where material parameter curves are smoothest (this typically excludes lower and higher limits of the operating frequency domain). Mathematically, a plot of TV shows the degree of fluctuation in the curves of refractive index and extinction coefficient for each separate thickness value.

Figure 3-15 shows an example of depth determination for a 1mm-thick silicon plate. The TV plot clearly shows the minimum that corresponds to the actual value of thickness (the smoothest curve). Figure 3-15B is a high resolution about the inflection in figure 3-15A. It estimates the actual thickness of the plate to be 1.0319(1) mm. This is one order of accuracy above the ability of standard micrometer gauges. Multiple micrometer gauge measurements give the thickness of the Si plate to be 1.031±0.001 mm. Note that the sample thickness found numerically corresponds to the effective thickness (taking into account surface roughness and non-uniformity), of the material in the area where the THz beam illumines the sample.

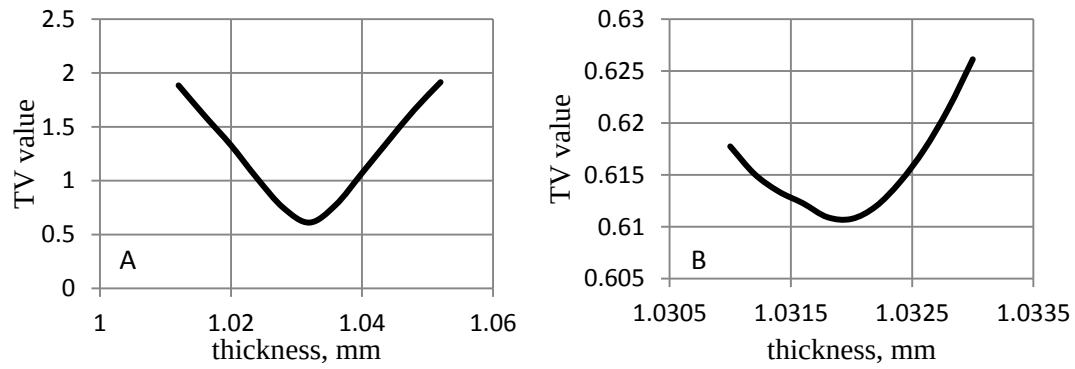


Figure 3-15. Total variation of material parameter curves, plotted for coarse (A) and denser (B) thickness ranges. The data is shown for a silicon plate of approximately 1 mm thickness. The same time-domain traces as in Fig. 5 with two distinct Fabry Perot (FP)-reflections have been used here. Note, that the first FP-reflection contributes most significantly to accurate determination of thickness.

Successful application of the TV method requires two important conditions to be met: FP-like ringing needs to exist and to be above the noise floor of the system. In fact, ringing of the THz pulse within the sample provides the physical means for accurate optical determination of the sample depth. If no internal reflections are detectable in the sample response, the TV algorithm does not work efficiently. In this case, instead of a sharp minimum in the TV curve, a broad, flat minimum extends for 10s of μm. Examination of equation (3.5) underscores this, where the FP term collapses to unity if no internal reflection is present. Spurious oscillations are consequently absent in parameter dispersion curves. The absence of optical ringing in the sample response can be for several reasons. In measurements of optically-thick samples employing a relatively short scan-length ($t_{scan} < 3nd/c$), application of n^{th} order transmission allied to the n^{th} order back-face reflection, arrive at the detector after scanning has finished. Interestingly, high absorption by a sample (e.g. $LiNbO_3$), does not lead to a significant decrease in the amplitude of FP-reflections. The feature most responsible for attenuating the n^{th} order internal reflection amplitude is scattering of THz radiation both within, and at the surfaces (i.e. impedance boundaries), of the sample and in the sample itself. It is principally this feature that limits the application of the TV methodology to study of bio-material samples like proteins, amino acids, sugars, etc. We have conducted a series of measurements on lysozyme, glycine and lactose pellets; and for all these there is no detectable ringing, while system dynamic range remains sufficiently high to distinguish THz pulse ringing within a 0.5 mm thick high-absorbing plate of lithium niobate.

Another significant factor influencing the performance of the TV method is sample alignment. The ideal case (as per the assumption of normal incidence in the use of Fresnel coefficients in (3.5)), is to set the sample-face-normal in-line with the THz beam axis. But this is not always possible, so a series of measurements have been performed to investigate the effect of face-normal deviation. Sample-tilt corresponds to an effective (secant) increase in sample thickness.

$$d_{new} = d_{actual}/\cos(\alpha) \tag{3.23}$$

Figure 3-16 shows the results of thickness determination of the tilted sample by the TV method and that calculated by (3.23). Sample thickness calculated by the TV method closely follows prediction (within the angle determination inaccuracy), proving the capability of the method to extract thickness correctly. For a sample tilt of 1˚ the sample depth changes by 0.3 μm as determined by both TV and equation 3.23. According to this analysis, in order to ensure thickness-error to be less than 0.1 μm due to sample misalignment, the sample normal needs to be set to within ±0.6˚ deviation off from the beam-axis.

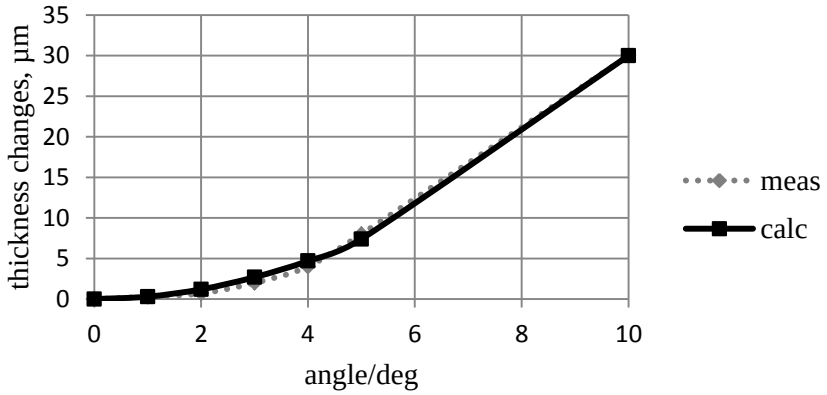


Figure 3-16. Changes in silicon plate thickness determined from measurements using the TV method, and from equation (3.23).

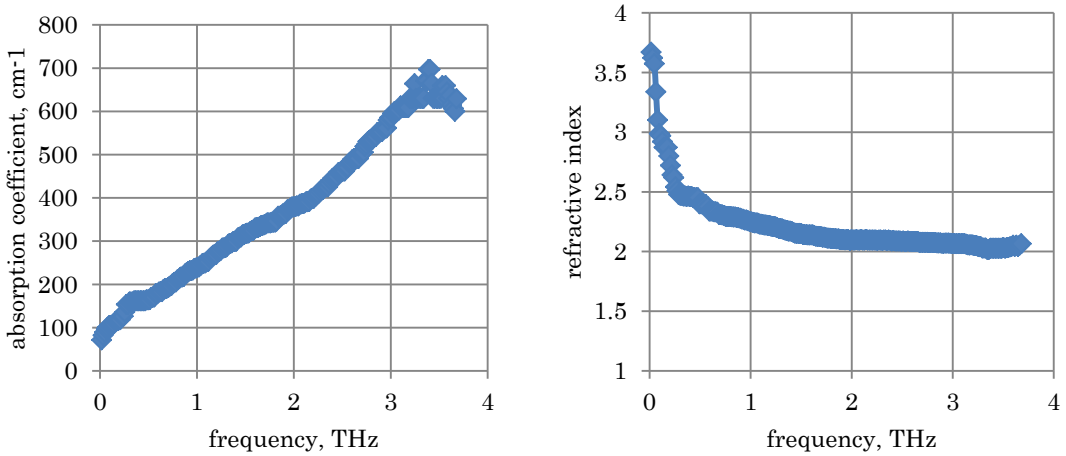
3.8 Precise material parameter extraction in multilayered structures

Another useful technique – extraction of material optical properties from multilayered structures – was employed form the paper of Wilk *et. al.* [Wilk'08]. The proposed algorithm allows one to effectively determine the optical properties of a layered structure. For example a liquid cell/sample holder (three layers) or thin films on a substrate. The algorithm implies that the properties of the windows of the sample-holding cell, or a substrate, are determined in advance, and there is only one material to

be characterized. The core of the method is also a complex transfer function which incorporates all possible reflections of a THz signal beam in each layer. As previously described, the theoretical TF is matched against the experimental one to result in an estimate of the complex refractive index of the sample. The algorithm was tested on our design of a liquid sample holder having TPX windows (Fig. 2-7), to extract the properties of a 100 μm thick water layer. TPX was chosen as the window material because of its transparency to THz radiation. Water, and any water-based solution, exhibits strong absorption of THz energy; therefore the thickness of the water layer has to be in the range of 100 μm. For thinner sample it is hard to control the thickness precisely and so the sample can be too thin for THz radiation to probe (via extended interaction), its properties. Water samples of even 200 μm lead to a dramatic dynamic range limitation and absorb the major portion of THz radiation available.

The extracted absorption coefficient and refractive index of a 100 μm water layer, spaced between two TPX plates are shown at Fig. 3-17 (upper plots). Our analysis estimates for the properties of water are consistent with published data [Bertie'96]. Above 3.5 THz the signal is below the noise floor for our current system setup. Successful estimation of water properties proves the possibility of doing concentration studies of a bio-molecule solution.

Fig. 3-18 shows the absolute errors associated with water properties extracted in fig. 3-17. The same notation of error terms is adopted here as in Fig. 3-9, with '1' – standing for error due to variance in time-domain response; '2' – being the error due to thickness uncertainty (± 2 μm); and '4' – shows the alignment error (± 2 degrees). All other sources of uncertainty in the water parameters are eliminated as a result of a) using the exact transfer function (eqn. 3.5) and b) the fact that the FP-reflections are negligibly small due to high water absorption. In contrast to silicon measurements, the main source of error for water is thickness uncertainty, since there are no means to accurately determine thickness numerically.



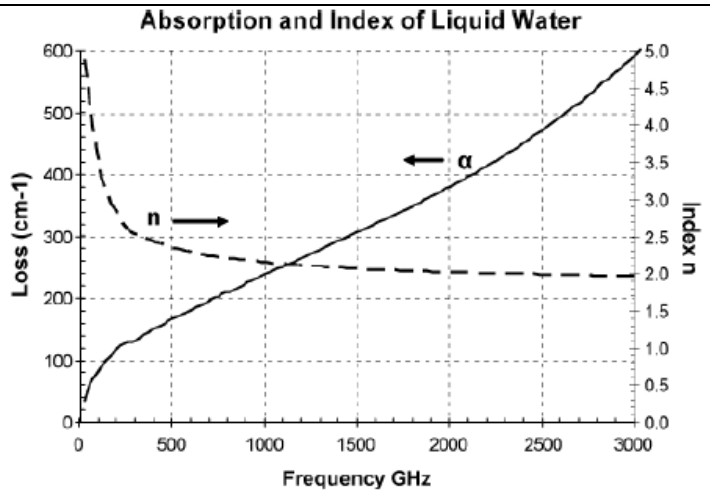


Figure 3-17. Upper plots – refractive index and absorption coefficient of water extracted from our measurements using a 3-layer extraction procedure. Lower plot: same properties of water taken from [Bertie'96].

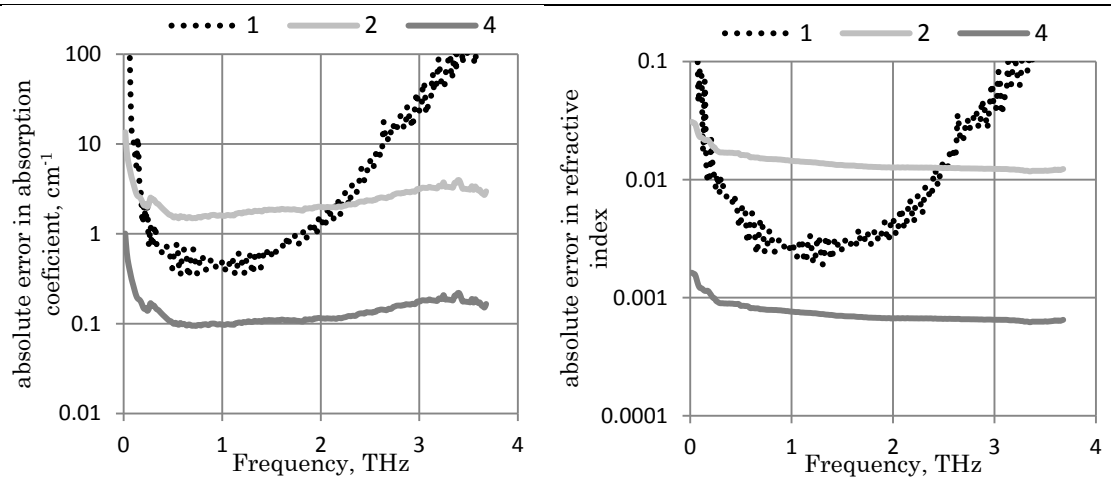


Figure 3-18. The absolute errors in absorption coefficient (left) and refractive index (right) of distilled water, estimated using the exact transfer function (3.5).

3.9 Brief summary for the chapter

The majority of groups active in THz spectrometry when undertaking an uncertainty analysis, if at all, primarily consider noise contributed from mechanical, optical and electronic components. Other sources of random and systematic errors introduced from various n and k estimation procedures are often not reported. Here, however, the effects of all sources of error are considered. The multiple internal reflections within the sample under test are found to be the dominant source of error if the advanced evaluating procedure is not used. A clear distinction has been set out between the best and worst case scenarios in terms of consideration of errors in determination of $\tilde{n}$ for silicon.

Different analysis procedures for estimation of n and k were also assessed. A study of the effect of supplying seed values for n and k in regression analysis routines was conducted and illustrated through physical measurements on high-purity silicon, quartz and highly dispersive lithium niobate. A guideline is consequently provided as to when certain parameter extraction procedures can be used and the affect that each has on the accuracy of estimating n and k.

A key emphasis of this chapter is the fact that relative errors of k are much higher than for n for silicon and low-absorbing materials. So choices on which extraction procedure to use should be carefully considered, especially when high-order reflection are present in time domain.

The main technical issues regarding THz-TDS parameter extraction and interpretation were discussed. Some experimental results are shown to illustrate the most important aspects concerning material characterization in THz domain. Discussion of the influence of alignment, sample preparation, scattering and humidity on the measured data was performed. Also material parameter extraction from multilayered structures was described.

The governing requirements have been discussed for accurate estimation of material optical properties by THz-TDS. The applicability and performance of numerical thickness determination using the total variation (TV) technique has been assessed. Accurate information on sample-depth, being the major source of uncertainty, is necessary for reliable material parameter (i.e. optical constants) determination. The accuracy of the TV method is shown to reach ±0.1 μm if FP-like multiple-order internal reflection response is present in the sample and alignment of the sample-normal is within ±0.6˚of the beam-axis.

Chapter 4. Terahertz Spectral Domain Computational Analysis of Hydration Shell of Proteins with Increasingly Complex Tertiary Structure

4.1 Overview

Water plays a key role in structural organization of bio-molecules and is a driving force that stimulates proteins to obtain their folded, functional state [Chaplin'06]. Over the last decade, studies of water solutions of different molecules in the THz and sub-THz domain have attracted growing interest by many [Matvejev'12, Arikawa'08, Leitner'08, Castro-Camus'08, Vinh'11, Heyden'10a, 12, Zhang'06, Kambara'10, Ebbinghaus'07]. They are stimulated by the fact that any molecule dissolved in water alters the dynamics of the surrounding water molecules to adopt quasi-coherent or organized character. This happens primarily via a re-organized, loose hydrogen-bond network. Much effort has been made to study the modified dynamics and dielectric properties of 'bio-water' (i.e. the solvation layer), by different experimental techniques such as: neutron scattering, nuclear magnetic resonance and x-ray crystallography, dielectric relaxation spectroscopy, etc. [Leitner'08] and by molecular dynamics simulations [Rocchi'98, Marchi'02, Sengupta'08,11,12a,12b, Sinha'08, Chakraborty'07, Pal'13, Bandyopadhyay'06, Xu'12, Ding'11, Heyden'10b,12b]. These techniques sense over a wide range of time-scales of bio-molecular dynamics, accordingly the probed-size of the hydration shell also varies. THz/far-IR time-domain spectroscopy (TDS) in particular possesses a unique ability to probe molecular-motion events of 100s of femtoseconds to 10s of picoseconds duration. The spectral vibration signatures of hydrogen bonds and of collective motion by globular proteins relevant to their function fall into this timing. THz radiation is also sensitive to coupling dynamics between protein and water and thereby serves as a probe of molecular structure and arrangement.

THz spectroscopy of water-based systems is challenging due to high absorption by water in the THz spectral domain. Thick samples therefore need to be reduced to 100s μm or less. The most recent approach is to treat a bio-molecular solution as a three-component system comprising protein, bulk water and hydration water [Ebbinghaus'07]. A complete frequency-resolved absorption for each component over the THz domain has not yet been established. A number of studies exist where THz absorption measurements were conducted as a function of a solution-concentration or

absorption of different functional states of proteins [Castro-Camus'08, Vinh'11]. It was shown that structural arrangements of ubiquitin and λ-repressor proteins (folded/partially-folded/unfolded) exhibit different absorption at THz energies [Heyden'10a]. These differences are thought to be attributed to changed hydration-water dynamics around the protein, as well as to structural flexibility of the protein. Concentration studies of protein solutions reveal non-linear changes in absorption coefficient [Heyden'10a]. This is usually explained by the onset of overlapping hydration shells of adjacent proteins. However the THz-absorption by hydration-water is still a point of controversy in the literature; some report its increased absorption to be related to bulk water [Heyden'10a, Ebbinghaus'07], others claim that hydration-water absorbs less than bulk [Arikawa'08, Zhang'06]. In part, these differences can be explained by the frequency domains or concentrations considered. An interesting THz study of myoglobin molar absorption was performed by Durbin et al. [Zhang'06]. According to their model, molar THz-absorption by myoglobin in solution is increased by more than one order of magnitude compared to a dry sample. The most rapid growth in absorption occurs towards the dilution limit. In another study by Heyden et al [Heyden'12a], protein is not considered as an absorbing particle in the THz domain, as its dynamics are much slower than the alternating THz field. THz absorption (0.1 – 3 THz, as accessible by most THz-TDS systems), of a protein solution is a complex phenomenon, where several key processes have to be considered. Firstly, the protein molar absorption due to its changed (compared to dry state) dynamics in a solution has to be taken into account. Secondly, the absorption coefficient of hydration water is different from bulk. And finally, accurate information on hydration shell extent into the bulk is needed, especially considering that the estimated depth varies with technique of determination. For instance, THz absorption measurements of the λ-repressor [Heyden'10a] suggest that a protein influences water dynamics out to 20 Å, while simulations are not able to detect any differences from bulk beyond 10 Å. While this probe-depth is more extended compared with NMR studies [Mattea'08, Qvist'09], that suggests that intracellular water that is not in direct contact with a bio-molecule, has essentially the same dynamics as bulk water. X-ray diffraction studies also reveal existence of 1-2 layers of water molecules around bio-molecule with properties different from bulk [Svergun'98]. Neutron scattering studies of hydration water can sense mainly the water molecules immediately adjacent to protein (not the extended hydration shell) and often at low temperatures [Frolich'08]. For instance, it has been shown that

hydrophobic amino acid side chains has a higher ordering of adjacent water molecules [Pertsemlidis'96].

Interpretation of THz absorption spectra of protein solutions has been facilitated by an increasing number of MD-based computational studies [Rocchi'98, Marchi'02, Sengupta'08,11,12a,12b, Sinha'08, Chakraborty'07, Pal'13, Bandyopadhyay'06, Xu'12, Ding'11, Heyden'10b,12b]. Molecular dynamics provide atomistic-level information on the interactions of molecules in solutions. Most of the studies reveal that solutes heterogeneously perturb surrounding water molecules, therefore water in the hydration shell shows anomalous dynamics compared to bulk water [Rocchi'98, Marchi'02, Sengupta'08, Sinha'08, Xu'12]. Rocchi et al. [Rocchi'98] have performed an MD simulation study of the dynamical properties of water at the plastocyanin interface. It unravels the sub-linear trend with time for the mean-square displacement (MSD) of hydration-water, resulting in anomalous diffusion. It was also shown that the rotational relaxation of water in the vicinity of lysozyme is 3 – 7 times slower than that in the bulk water, depending on the definition of hydration shell in the calculation [Marchi'02]. Sengupta et al. [Sengupta'08] investigated the hydration dynamics in a partially denatured human α-lactalbumin and discovered that denatured conformers are less uniformly solvated and have an enhanced water molecule dynamics. A considerable amount of simulation studies on the hydration of proteins have been done by the Group of Bandyopadhyay [Sengupta'08,11,12a,12b, Sinha'08, Chakraborty'07, Pal'13, Bandyopadhyay'06]. For instance, the thickness of the hydration shell of the villin headpiece sub-domain HP-36 was investigated by different methods including MSD, re-orientational correlation function analysis, H-bond time correlation function analysis and velocity autocorrelation function analysis of water molecules. Simulation results demonstrated that an heterogeneous influence of different helical segments on the dynamics of water around them is limited to the first hydration layer [Sinha'08]. Similar results have been obtained by Xu et al [Xu'12], where it was shown that power spectra (VDOS) for the water molecules hydrogen-bonded to different planes of the antifreeze protein, exhibit distinct spectra in the 1–4 THz spectral domain. In another study the vibrational spectrum of water in the hydration shell has been addressed [Chakraborty'07]. There it was shown that the O-O-O bending mode experiences a clear blue-shift in the first hydration layer, which is especially pronounced for water molecules hydrogen-bonded to protein. It was reported that structural flexibility of protein can be also studied. It was found that the overall flexibility of lysozyme is primarily controlled by a few large-amplitude bi-stable motions exhibited by two coils

[Sinha'11]. Ding et al. have used VDOS to analyze the contribution and make assignment of different structural elements of alanine-rich peptides to vibrational bands observed in the experimental spectrum [Ding'11]. The most rigorous approach to calculate the THz spectrum of water is quantum mechanically. Extensive study of water from first principles, resolved in time and space, have been performed by Heyden et al [Heyden'10b]. Their findings indicate that the contribution to spectral intensities around 2.4 THz is dominated by group motion of H-bonded molecules within the second solvation shell, indicating also the presence of non-negligible, third-shell effects. The topic was further developed by Heyden et al [Heyden'12b], showing that non-polarizable water models, as employed in this chapter, are very capable in reproducing the low-frequency, intermolecular vibrations of water, since the static molecular dipoles dominates over electronic polarization. However there is no *systematic* study concerning the depth of the hydration shell, especially as dependent on protein size. My methodology to estimate the hydration shell depth is based on its component contribution to the THz vibrational spectrum, i.e. as based on the VDOS. My calculations focus mainly on THz-domain spectral properties of solutions.

4.2 Gromacs simulation details

All molecular dynamics simulations were performed using the Gromacs package [Pronk'13]. It is based on numerically integrating Newton's equations of motion for atoms at discrete time-steps. The Amber99 force field was applied to characterize molecular interactions for all bio-molecules studied. Each protein was then solvated in a cubic cell with a TIP3P water-model, leaving 1 nm gap between the protein to the edges of the box. Boxes respectively contained 2239, 2683, 7026 and 11713 water molecules for TRP tail, TRP-cage, BPTI and lysozyme. To neutralize the non-zero charge of protein, chlorine ions were used (1 for TRP tail and TRP-cage proteins, 6 – for BPTI and 8 – for lysozyme). Energy minimization was performed using a steepest-descent algorithm to avoid any poor contacts between protein and water molecules. Further, a 100 ps equilibration was run to allow water molecules to fully envelope the protein. During such runs a Berendsen temperature-coupling was used to rescale the atomic velocities and let the solutions reach 300 K. A Parrinello-Rahman pressure coupling was simultaneously adopted to stabilize the system pressure at 1 atm. Electrostatic interactions were treated by the Particle-mesh Ewald algorithm, employing a 1 nm cut-off radius. Three dimensional periodic boundary conditions were applied. Pressure and temperature in the system were monitored at each step to ensure convergence towards

set-point values. The resulting system configuration was used to start 200 ps production runs employing the same parameter settings. An integration time-step of 2 fs was used while atomic positions and velocities were recorded at every second step. Bulk (i.e. pure) water properties were extracted from the simulation of 1026 water molecules, also employing the above same parameter settings.

The contribution of each hydration-water layer to the vibrational spectrum is estimated based on a normalized VDOS, i.e. the Fourier transform (FT) of the ACF of atomic velocities:

$$\mathrm{VDOS(f)} = \int \frac{\langle \vec{v}(0)\vec{v}(t) \rangle}{\langle \vec{v}(0)\vec{v}(0) \rangle} \exp(i2\pi ft)dt. \qquad (4.1)$$

The normalized velocity ACF of the selected atoms is calculated using the post-processing tools of Gromacs. The motion of the atoms is of an oscillatory nature (atomic velocities self-correlate in a periodic manner), so the VDOS characterizes the oscillation as an energy-band in the frequency domain. The VDOS spectrum shows the vibrational bands of covalently-bonded and H-bonded atoms (for instance O-O-O bending, O-O stretching, O-H stretching), including anharmonic effects. Note, that the vibrational spectrum estimated by VDOS does not directly correspond to the experimentally measured absorption, which is calculated as an integral of the total dipole ACF and requires a quantum dynamical approach [Heyden'10b]. Although the VDOS does not determine the exact molar absorption of the solution, it shows the relative intensity and spectral shifts in the vibrational peak positions. By comparing the VDOS of water molecules in successive layers around a protein to that of bulk water, we determine the size (or depth), of the hydration shell. The VDOS curves were calculated for stratified 3Å layers of water molecules (0-3 Å, 3-6 Å, 6-9 Å, 9-12 Å) over 10 ps intervals over the whole trajectory. These integral characteristics of each layer are in fact slightly different from instantaneous values due to molecular diffusion. However the maximum distance any molecule travels from the layer for 10 ps is less than 1 Å; moreover molecules diffuse in both directions evenly (i.e. both towards and away from the protein), which to a significant extent is mutually compensatory and lessens the error caused by diffusion. All VDOS curves in figures below are averaged over 10 trajectories. One of the useful outcomes of calculating the VDOS spectrum is access to the diffusion coefficient (D), evaluated by the Green-Kubo relation [Liu'04]:

$$\mathrm{D} = \frac{1}{3} \int \frac{\langle \vec{v}(0)\vec{v}(t) \rangle}{\langle \vec{v}(0)\vec{v}(0) \rangle} \mathrm{dt}. \qquad (4.2)$$

However, calculation of diffusion coefficients for interfaces, or inhomogeneous regions like the hydration shell, proves to be problematic [Liu'04]. The appropriate

corrections for limited box-size and simulation time have not been taken into account here. Therefore the value of *D* is referred to as a perturbation coefficient. It reflects a measure of disorder in water dynamics caused by protein. Its values are proportional to the actual diffusion coefficient and are used later as an additional criterion to determine the depth of the hydration shell.

4.3 The structure of selected proteins

Four different bio-molecules, with increasingly complex tertiary structure (corresponding to an increased number of secondary structure elements), were selected for the analysis of their hydration shell, namely: TRP-cage$_{13\text{-}20}$ peptide, TRP-cage, BPTI and lysozyme. Proteins were chosen purely based on an increased structural complexity, starting from a single peptide chain to lysozyme, a globular protein. Both molecular weight and residue numbers scale up approximately 2-3 times successively larger proteins, starting from TRP-tail. An increasing variety of tertiary structure of proteins can be seen in Fig. 4-1 (D-A). BPTI contains fewer helical elements than TRP-cage protein, however a prolonged β-sheet and peptide chain provide the required structural complexity. Brief information on the secondary structure of these is given in Table 4-1 and shown in Figure 4-1. PDB files of lysozyme and BPTI do not contain disulfide bridges, which have only a minor effect on the present study.

Table 4-1. Description of considered proteins

Bio-molecule/property	Molecular weight, g/mole	Residue number (amino acids)	Secondary structure elements
Lysozyme (PDB: 2LYZ)	14313,3	129	16 helix turns + 2 β-sheets
BPTI (PDB: 4PTI)	6517,6	58	2 helix turns + 1 β-sheet
TRP-cage (PDB: 1L2Y)	2169,4	20	4 helix turns
TRP-cage tail (13 to 20 residues)	784,9	8	none

Figure 4-1. 3D structure of (from left to right): (A) lysozyme, (B) BPTI, (C) TRP-cage and (D) TRP_{13-20} tail. Helices are shown in pink, β-sheets – in yellow, peptide chains with no secondary structure elements – in grey.

4.4 Hydration shell: buried vs surface water molecules

The aim here is to investigate how dynamically different the hydration shell of each of the chosen proteins is, and what is the depth of the hydration shell. Figure 4-2 shows the velocity ACF and VDOS curves calculated separately for oxygen and hydrogen atoms in bulk water and in the first hydration layer of lysozyme (contains all water molecules within 3 Å of protein atoms). As expected, the dynamics of hydrogen atoms becomes uncorrelated at approximately 0.15 ps and is much faster that oxygen dynamics, having a correlation time of 1 ps. Both curves for oxygen VACF have a small extremum (bump) after initial descent which is a characteristic signature of the 'caging' effect (i.e. where oxygen molecules collide with neighbour-molecules, enveloping them). For oxygen atoms of the first solvating layer, this bump is smaller, indicating their more restricted motion. The VACF of hydrogen atoms in the first hydration layer show only minor differences from bulk-water hydrogen, which is reflected in a slight decrease of their perturbation coefficient. VDOS curves were obtained by integrating VACF only in the correlated domain. Bulk-water oxygen exhibits a well-defined band centered at 1.1 THz. This peak is attributed to the bending motion of triplets of H-bonded oxygen atoms [Sinha'08]. It has a strong Raman band but is a quiet IR absorption mode. For water molecules bound to protein, the peak is blue-shifted by 0.4 THz and its amplitude lowered. The lower peak amplitude indicates the vibrational mode to be less pronounced due to influence of the protein. The frequency shift corresponds to the strengthening of the H-bonds between oxygen atoms in the solvating layer. In fact the peak also broadens by 9%, which means that this mode became less harmonic. Note that above 2.4 THz, hydration water exhibit a larger VDOS compared with bulk. While the VDOS of hydrogen atoms is rather flat in THz domain and changes in it due to protein proximity are much less pronounced. The dynamics of oxygen atoms are therefore seen to chiefly govern the THz vibrational spectrum of

hydration water. For this reason we study the properties of the solvating water shell based on the VDOS of oxygen atoms only. This assumption is valid only in THz frequency domain (below 10 THz).

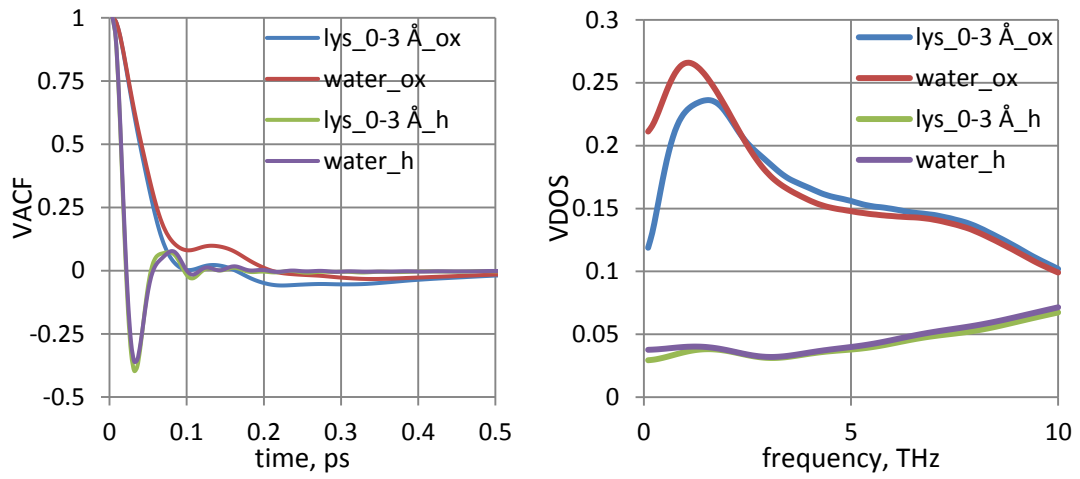


Figure 4-2. (Left) Comparison of the VACF and (Right) VDOS of oxygen and hydrogen atoms for bulk water and for the first 3 Å solvating layer of lysozyme containing 353 water molecules. 'ox' stands for oxygen atoms, 'h' – for hydrogen atoms.

Solvation dynamics have been studied for only a limited number of molecules, primarily small and fast-folding proteins [Lindorff-Larsen'11, Kubelka'04]. Mainly single proteins were considered, and a systematic comparative analysis of solvation dynamics for the different complexity of protein tertiary structure is not available.

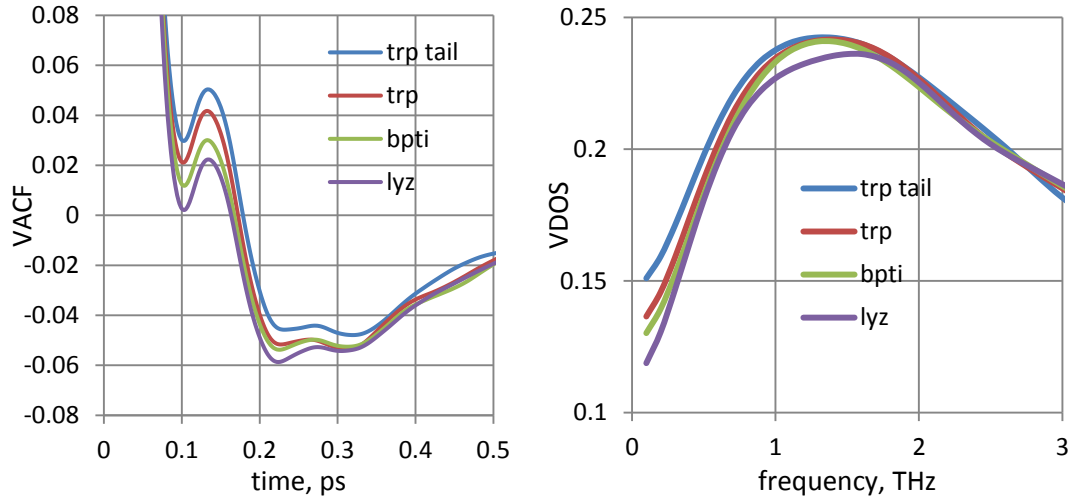


Figure 4-3. (Left) Velocity auto-correlation function and, (right) vibrational density of states of oxygen atoms in the first 3 Å hydration layer around associated proteins

Figure 4-3 depicts the VACF and VDOS of water oxygen atoms in the 3Å shell within the four bio-molecules selected for study. A clear trend is observed of oxygen

atoms being more constrained at the interface of the protein with more complex tertiary structure. The difference between curves is well above the error which is approximately the same for each trace and is equal to 10^{-4}, estimated from the standard deviation of the VACF over 10 simulations. For bio-water in close proximity to the protein, these results show different features in the vibrational spectrum of the solvating layer in accord with the diversity of 3D protein-shape. Note that the perturbation coefficient has dropped in line with the complexity of the tertiary structure.

Note also that the 3 Å hydration layer considered above *includes* the molecules buried in the protein-interior as well as those immediately bound to the outer-surface of the protein. To unravel the physics behind the systematic differences in the first hydration layer among proteins we separated the contributions from those waters buried deep within the protein-interior and those solvating the protein surface. An envelope defined by the locus of points 3 Å normally-distant from the protein surface, forms a boundary for deciding which waters are surface and which are buried. Positive distance (d) from the boundary is directed back into the volume of the envelope. Surface waters are defined by $0 < d \leq 3$ Å; buried waters are $d > 3$ Å. The results of this analysis are shown in figure 4-4. It depicts the VACF and VDOS of the buried and surface waters for the four considered bio-molecules.

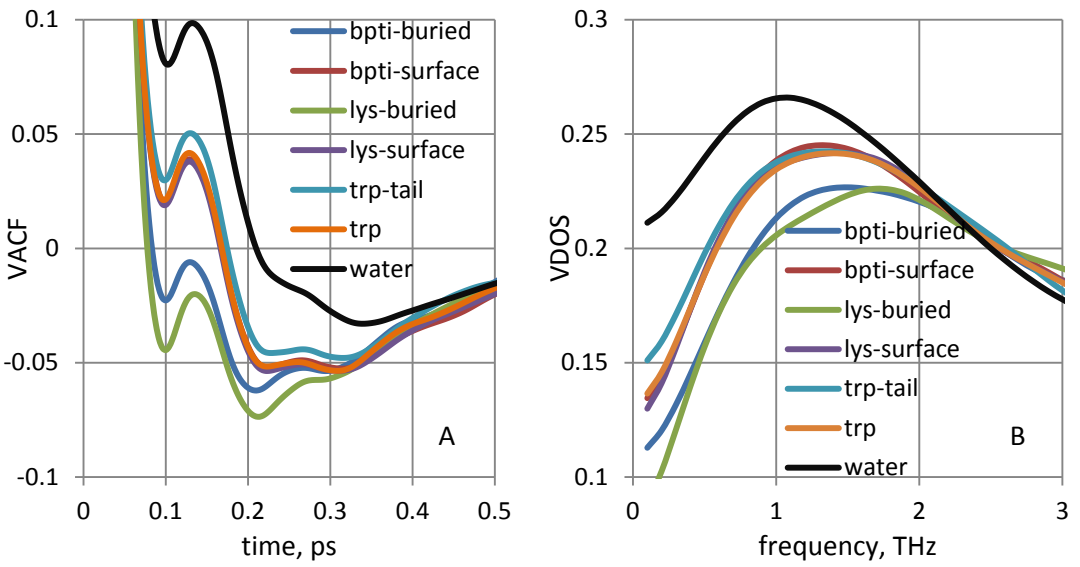


Figure 4-4. (A) Velocity auto-correlation function and (B) vibrational density of states of buried and on-surface water molecules for the bio-molecules studied. There are no buried water molecules for TRP-cage and TRP-cage$_{13\text{-}20}$ tail (all molecules solvate the surface). The corresponding properties of bulk water are also shown for comparison. Note, that in both plots, curves for surface water molecules of TRP-cage, BPTI and lysozyme closely overlap.

The difference in solvation dynamics for the proteins studied is clearly shown to originate primarily from highly-constrained buried water molecules (figure 4-4). The caging effect appeared to be more pronounced for buried molecules (figure 4-4A), resulting in a more pronounced frequency-shift of the O-O-O bending mode (figure 4-4B). Due to the simple structure of the TRP-cage and TRP-cage$_{13\text{-}20}$ tail, their solvation layers consist only of 'surface' molecules. Water molecules bound to each of the proteins' surfaces show similar dynamical behavior. Only the solvation-layer dynamics of the TRP-cage tail differs due to higher mobility of water molecules bonded to the peptide, caused by the absence of crevices and clefts on its surface. These findings are key for understanding the dynamics of protein hydration. It is safe to predict that similar effects take place for proteins of arbitrary size, since the relative numbers of hydrophilic and hydrophobic residues are comparable for most of globular proteins. However this assumption needs additional verification by simulations. The average number of buried molecules is 95 and 41 for lysozyme and BPTI respectively. These water molecules, tightly bound to protein, can be regarded as a part of the structure of the protein, and accompany its global motions. Therefore, the contribution of these molecules to the THz absorption spectrum is much weaker than that of bulk water [Vinh'11].

Hydration dynamics is better understood if one interprets consecutive layers of the hydration shell as 'matching' layers between relatively slow protein motion and fast water dynamics. The layers matching protein and bulk water dynamics have less retarded motion compared to the layers between protein and another layer of constrained waters. In this later group, molecules are confined within more than one protein residue, or are trapped in the interior of the protein and accordingly have slower dynamics. It is worth noting that hydration-water beyond the first shell exhibits minor differences in dynamical behavior with the proteins selected here for study.

The VDOS of water oxygen atoms in stratified layers of 3 Å thickness (which is approximately equal to a single-molecule layer) within the lysozyme protein is plotted in Figure 4-5. The trend of steady increase in perturbation coefficient with distance from the protein is clearly observed here. The presence of protein blue-shifts the O-O-O bending peak, while beyond 6Å it regains the spectral position for bulk-like response. The same spatial extent of 6 Å is observed for the hydration water, having an increased spectral intensity compared to the bulk above 2.4 THz. The same effect is present among the other proteins, though less pronounced. This characteristic frequency shifts slightly lower for smaller proteins, reaching 2.1 THz for TRP-cage tail peptide.

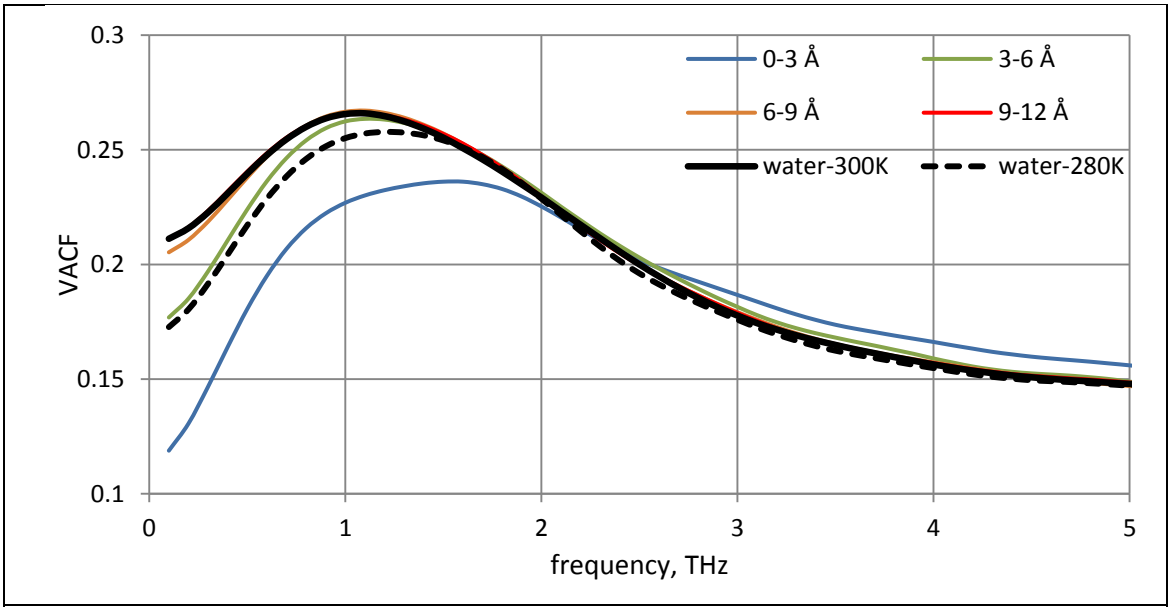


Figure 4-5. The vibrational density of states of oxygen atoms in water at 300 K and 280 K and in successive hydration-shell bands of 3 Å thickness each for lysozyme.

Additionally, the similarity has been analyzed between retarded-water dynamics in the protein shell and water dynamics resulting from temperature decrease. The cubic water box, containing 1026 water molecules, is simulated at 280 K with other conditions identical to previous runs for bulk water at 300 K. The nature of retardation of water that is solvating a protein is found to be different to the low-temperature water dynamics. Similar behavior is seen up to a characteristic frequency, with spectral peak-position blue-shifted and attenuated. Above this frequency, hydration-shell water has a higher VDOS, and 280 K water has a lower VDOS compared to bulk water at 300 K. Such behavior is characteristic of the first and partially of the second hydration layers, but not of a third. So an increased VDOS of hydration-shell over bulk-water, above characteristic frequency, can be thought of as the specific signature of protein-water H-bond dynamics. This may be a result of the coupling with protein, having a typically high density of normal modes in this range. This behaviour is claimed to interpret some features of THz absorption spectra in protein solutions. For instance, it has been related to an increased absorption by ubiquitin solution over the buffer, above 70 cm^{-1} (2.1 THz) [Heyden'10a]. Similar features of VDOS have been shown for the hydration shell of villin headpiece protein by Chakraborty et al [Chakraborty'07].

4.5 Determination of hydration shell size from MD simulations

The estimation of hydration shell depth has been a question of debate recently. In this computational study we have estimated the size of a hydration shell based on its spectral

characteristics. Two complementary methods are used: one is investigation of the evolution of the perturbation coefficient with increasing distance from a protein; the other being the total difference of the VDOS spectrum between bulk and hydration water, evaluated as a sum of absolute differences at each frequency point over the 0-5 THz spectral domain. These two sets of curves are plotted against distance from the protein in figure 4-6.

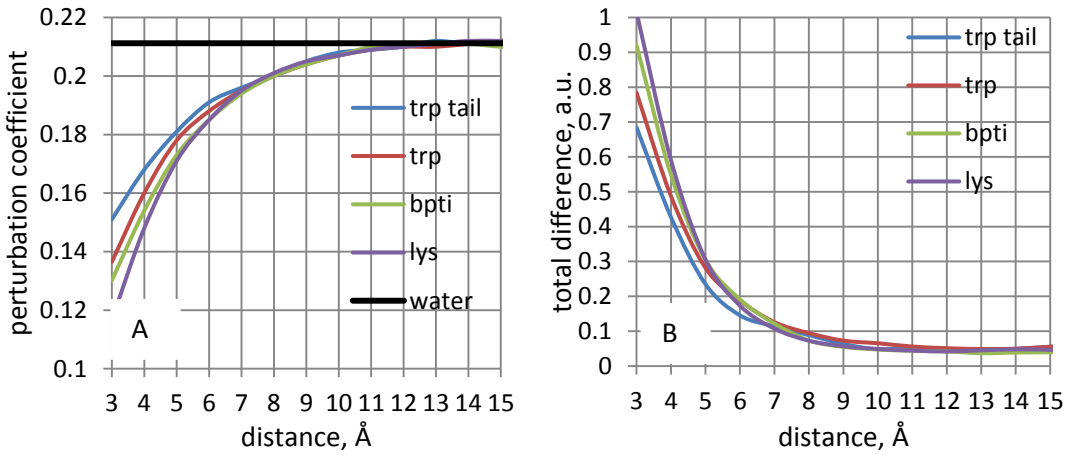


Figure 4-6. Plots of perturbation coefficient for the four selected proteins (A), and total difference in vibrational density of states between the hydration layers and water over 0-5 THz spectral domain (B), depending on the distance from the protein surface. Both plots were calculated for a successive overlapping 3 Å layer thicknesses, i.e. 0-3 Å, 2-5 Å, 3-6 Å … 13-16 Å. Note: buried molecules were also included in the calculation of properties of hydration shells.

Both approaches suggest that the dynamical hydration shell, whose spectral properties differ from bulk, extends out to 10 Å for all bio-molecules considered. Since the differences in surface water dynamics are negligible for the proteins and a $TRP_{13\text{-}20}$ peptide, the hydration shell size is approximately the same for all. Note, that surface water molecules have slightly faster dynamics for $TRP_{13\text{-}20}$ peptide (figure 4-4). This however does not affect the depth of the hydration shell in a measurable way. Interestingly, the total difference curves are more flat towards the end of the hydration shell compared to perturbation coefficient curves. This feature is explained by the fact that for hydration water beyond 6 Å the difference in VDOS compared to bulk water is pronounced only at low-THz frequencies (see figure 4-5, 6-9 Å curve) and is therefore well captured by the perturbation coefficient, but not by the total difference, where this low-frequency distinction is averaged (or diluted) over the whole 0-5 THz spectral domain.

The main property of interest however is the total integral characteristics of the whole hydration shell. The perturbation coefficient of oxygen atoms, averaged over the whole hydration shell, is lower than that of water respectively by 9%, 12%, 14% and 16% for TRP-cage tail, TRP-cage, BPTI and lysozyme. Furthermore, the overall shape of the hydration shell VDOS is similar to that of the first hydration layer (figure 4-6), with the characteristic frequency in the range of 2.1 – 2.4 THz. Based on the distinction in the perturbation coefficient, it is save to predict that the absorption coefficient of the hydration-shell water molecules is different from bulk water molecules and varies in relation to the shape of a protein.

4.6 Concluding remarks

Water in the hydration shell of proteins with different tertiary structure will respond to THz radiation with characteristic differences. According to molecular dynamics simulations, hydration shell thickness, as estimated by its total difference in VDOS and by its perturbation coefficient of water solvating protein, is shown to extend out to 10Å from the surface of a protein and does not depend on the size of a protein. Furthermore the integral perturbation coefficient of the whole solvation layer is found to be increased for larger proteins due to a higher retardation rate of water molecules in their shells. The THz vibrational signature of the solution depends on the amount of highly-retarded water molecules buried in the interior of a protein. The immediately-bound water molecules to the surface of the proteins were shown to have similar dynamical properties for the selected three proteins and a TRP-cage$_{13\text{-}20}$ peptide. However, the number and degree of constraint of internal (i.e. buried) waters, proved to be the dominating source of differences in solvation dynamics. The hydration shell of larger proteins tends to have lower vibrational density than smaller proteins below a characteristic frequency (2.4 THz for lysozyme), and higher above. These computational studies are to be verified by corresponding experimental THz time-domain spectrometry. Note particularly, that the VDOS vibrational spectrum cannot be directly related to THz-TDS that probes dipole moment fluctuations. However, the dependence of the characteristic frequency on protein size, as found from the VDOS (2.1, 2.25, 2.35, and 2.4 THz), for TRP-tail, TRP-cage, BPTI and lysozyme respectively, can be verified with corresponding THz absorption measurements.

Chapter 5. Investigation of hydration shell thickness and electromagnetic field absorbance of solvated protein molecules in the sub-THz frequency domain.

5.1 Introduction

In their native form proteins exist and perform their function in a solvated state [Chaplin'06]. Protein solutions have been studied by many experimental techniques and over a broad frequency domain from 10s of GHz up to few THz [Saha'12, Laurette'10, Leitner'08]. Compared to established FIR spectroscopies and NMR, THz (and sub-THz spectrometry, particularly), of protein solutions, is still in its infancy, but provides useful information of picosecond to sub-nanosecond process-dynamics of solutes [Heyden'12a]. THz spectrometry is commonly performed as Time Domain Spectrometry (TDS), driven by femto-second pulse-width lasers [Exeter'89]. Such ultra-short pulse-duration, constitutes an effective probe of collective vibrational modes of bio-molecules, long-range mutual coupling between solvent and solute and the bending and stretching of hydrogen bonds [Ebbinghaus'07, Heyden'12a]. An extensive review on the application of THz spectrometric analysis of bio-molecules has been published by Falconer and Markelz [Falconer'12]. It discusses the current progress in acceptance of THz radiation as a useful analysis tool in the biochemistry community and highlights the need for validation of THz spectroscopic results against an orthogonal experimental approach. The frequency domain of these studies spans from 70 GHz to 3 THz and is limited mainly by the dynamic range of a given THz-TDS system and increasingly high water absorption at higher THz frequencies. High absorption of THz radiation by water initially limited the examination of solutions, resulting in THz radiation being mainly utilized for characterization of dry-state bio-molecules [Jin'10; Whitmire'03; Markelz'00]. With THz sources and detectors becoming gradually more available and efficient, THz spectrometry promises to turn into a standard tool in biochemistry. Despite a large number of studies on hydration dynamics of bio-molecules the exact picture is still unclear. The trends in THz absorption of bio-solutions have not been studied consistently over the whole THz spectral domain and with respect to protein size and structure. Ideally, one would seek to obtain the complete frequency, concentration, and temperature-dependent information about the given complex dielectric properties of a solution. Often only part

of the above information is recorded and analyzed to provide some insight into bio-molecular solvation dynamics. The dynamics of constituent components in bio-solutions is attracting growing interest but is still not well understood, results even being sometimes highly controversial. For example BSA solution in water at 0.1% w/w concentration was shown clearly to have higher absorption compared to water in the 2-2.5 THz spectral domain [Saha'12] (as judged by the higher imaginary part of dielectric constant), while the absorption of BSA in 50mM phosphate buffer at 101 mg/ml (effectively the same concentration as in [Saha'12]), was shown to have less absorption than the buffer by approximately 8% over the frequency domain from 0.5 to 2.5 THz [George'08]. The concentrations in these two studies differ by only 1%, but it is unclear whether a rather dilute phosphate buffer could cause such disparate results. These findings also do not completely agree with those of M. Heyden and coworkers [Heyden'10a], where the concentration-dependent study on the λ-repressor has been performed at 2.4 THz. The highest THz excess was detected at 6.5 mg/ml. At 13 mg/ml the absorption of the solution was the same as that of the buffer and decreasing. Even considering that the λ-repressor has a molecular weight eight times smaller than BSA, it is counter-intuitive that a 0.1 w/w BSA solution would exhibit excess absorption at 2.4 THz. Importantly, none of the above studies have considered the influence of the intermolecular stretching mode in water, centered at 5.3 THz. It was shown [Yada'08] that energetic contribution from this mode constitutes third of the imaginary part of the dielectric constant at 2 THz.

Increasing numbers of research groups around the world have studied bio-molecules in a solvated state by means of THz radiation. Significant progress has been made in interpretation of complex dielectric spectra of solvated bio-molecules partially facilitated by molecular dynamics (MD) simulations [Gekle'12] and technological advances in THz spectrometry [Jepsen'11]. Studies by the Group of Havenith are focused mainly about 2.1-2.7 THz, with the probing radiation produced by a p-germanium laser, and applied to the study of a wide range of bio-molecules e.g. DNA, proteins, peptides, amino acid and carbohydrates. Their findings show that the $\lambda_{6\text{-}85}$-repressorand solutions of ubiquitin proteins exhibit THz excess at specific low concentrations (of the order of 10mg/ml) [Heyden'10a]. Such behaviour is claimed to be caused by an increased absorption of hydration-water compared to bulk-water. These same studies also predict extended shell-size of up to 20 Å derived from the concentration of peak THz excess [Ebbinghaus'07] and not detectable by other techniques. Even more distant interaction between protein and solvent reaching 20-40 Å

has been predicted by Heyden et al. [Heyden'12a] based on the cross-correlation of solute and water dipoles. The possibility of peptides to perturb water dynamics beyond 10 Å was also claimed by Ding et al [Ding'10], where the authors experimentally derived the hydration-shell size of alanin-rich peptides to be 11 to 17 Å. Another experimental study of nucleotides in solution proved hydration shells to extend out to the fourth layer of water molecules (≈ 10 Å) [Glancy'10]. The temperature-dependence of THz excess has been monitored both for the $\lambda_{6\text{-}85}$-repressor [Ebbinghaus'07] and antifreeze glycoprotein (AFGP) [Ebbinghaus'10]. The THz excess peak for AFGP shifts towards lower concentration with decreasing temperature, indicating a more extended hydration shell; and retarded water dynamics in this larger shell does not favour freezing. It was also shown that the frequency domain of 2.1-2.8 THz is sensitive to both the functional state of protein and the amino acid sequence, with the native protein solutions showing the highest THz excess [Ebbinghaus'08]. At the moment the only group (to the best of my knowledge), that reported an increased THz absorption by water in the hydration shell of proteins, compared to bulk-water, at around 2.5 THz is the Group of Havenith [Heyden'10a]. The same study contains the only piece of evidence of both THz excess and defect recorded in the same measurement run (for ubiquitin solution showing a THz excess above 2.15 THz, and THz defect below). A number of studies [Arikawa'08, Zhang'06] have shown that a protein hydration-shell absorbs less than bulk-water in the frequency domain below 1.5 THz.

MD simulations often assist in interpreting experimental spectra. For instance, experimental studies were complemented by appropriate MD simulations [Heyden'10a] that relate THz excess of proteins to an increased vibrational density of states of hydration-water over that of bulk-water in the frequency domain above 55 cm^{-1}. Also in a combined experimental and MD study of aqueous peptides Ding et al. [Ding'11] used vibrational density of states to assign and analyze contributions of different structural elements to the absorption spectrum. The retarded dynamics of water molecules in a hydration-shell of protein [Heyden'10a, Niehus'11] was also studied by the autocorrelation function (ACF) of H-bonds, rotational ACF and mean square displacement (MSD). Generally in MD simulations of the protein hydration shell no differences between hydration-water and bulk were found beyond 10 Å.

Along with proteins, the 2.1-2.7 THz absorption of amino acids and homogenous mono-, di-, and tri-peptides was investigated [Niehues'11]. At first the dynamics of water molecules were analyzed that were solvating model hydrophobic and hydrophilic particles. It was then shown that the slope of the plot of THz absorption

versus frequency is directly correlated with the hydrophobicity scale of amino acids, based on the non-polar accessible surface. Studies of carbohydrate solutions [Leitner'08, Heyden'08] reveal a non-linear THz absorption attributed to overlapping hydration shells of neighbouring molecules. Based on three-component absorption model (bio-molecule, bulk-water, hydration-water), the hydration-shells were determined to extend out to 3.7 Å, 5.7 Å, 6.5 Å for glucose, lactose and trehalose, respectively. One of the key findings was that the absorption coefficient of hydration-water was on average 10 cm^{-1} higher compared to bulk-water in the frequency domain 75-95 cm^{-1}. This is explained in terms of coherent oscillations of hydration-water and solute [Heugen'06]. In agreement with Leitner et al [Leitner'08], Arikawa and coworkers [Arikawa'08] have found the hydration shell of sucrose to extend out to 7.4 Å away from the solute. Absorption of water molecules in the hydration-shell of alcohols has been addressed by Matvejev et al [Matvejev'12]. They found that water absorption by the hydration-shell is on average 0.774 times that of bulk-water absorption at 0.28 THz. This is in contrast to the hydration-shell of saccharides (lactose, trehalose, glucose) that have higher absorption than bulk-water [Heyden'08] due to a higher density of hydrophilic groups in the molecule, resulting in the mobility of hydration-shell waters.

The absorption spectrum of solvated bio-molecules over the THz spectral domain, as well as techniques to interpret it, has been extensively investigated by the Group of A. Markelz [Chen'07, Knab'07, He'11]. It was shown that THz radiation is sensitive to solution phase binding as demonstrated on triacetylglucosamine binding to lysozyme [Chen'07, Knab'07]. The recorded THz spectrum showed a clear decrease in absorption with binding. The influence of hydration levels on protein dynamics was also addressed [Knab'06]. A well-defined dynamical transition, revealed by a more rapid rise in complex refractive index, was observed at hydration levels when the first hydration shell is filled [Knab'06] or at hydration levels of approximately 0.27 grams of water per gram of bio-molecule [Kambara'10, Vinh'11]. Similarly, bio-molecules have been shown to have a dynamical transition at around 200 K [Yamamoto'12; Markelz'07], related to the onset of anharmonicity and is detectable in THz spectra. Another study [He'11] reveals the existence of structural collective modes for cytochrome *c* protein. These were confirmed by analyzing absorption spectra with quasi-harmonic vibrational modes' density and the dipole-dipole ACF.

The collective dynamics of lysozyme [Xu'06a] and BSA [Xu'06b] were investigated by Xu et al. using THz absorption spectrometry over a wide frequency

domain of 0.075-3.72 THz. It was shown that molar absorption of these proteins increases monotonically for BSA and saturates above 2 THz for lysozyme. It was found that molar absorption of BSA does not significantly depend on concentration, except for very dilute solutions where an uncertain growth of molar extinction was observed [Xu'06b]. A sharp increase in molar absorption towards the dilution limit was also detected for myoglobin protein [Zhang'06]. At concentrations of 70-98 (wt%), molar absorption of myoglobin was reported to be more than one order of magnitude higher compared to the dry state.

Vinh and co-workers have performed a highly sensitive vector network analyzer-based spectroscopy on concentrated lysozyme solutions covering 65 to 700 GHz [Vinh'11]. They discovered the presence of approximately 165±15 water molecules tightly bound to protein and behaving like an integral part of it. In the modeling of these results the authors introduced a 250 GHz cutoff frequency below which the density of vibrational modes is zero. This has significantly improved agreement of experimental and modeled dielectric spectra. The same proteins, along with alcohols, have been also studied at different concentrations by micro-fluidic system [Laurette'10,12] over 50-110 GHz. Such a system is based on a coplanar-waveguide and is propounded to be a sensitive (5 mg/ml) tool for biological liquid metrology.

A number of different approaches to the interpretation of recorded THz spectra have been used. Various studies sometimes do not agree in details, some missing the detailed description of experimental procedures and materials used. There exists, therefore, a need for additional systematic studies of different bio-molecules in a solvated state.

5.2 Methods and simulation procedures

The following proteins, ranging from low to high molecular weight, were chosen for this study: lysozyme, myoglobin and bovine serum albumin (BSA). All proteins were purchased from Sigma Aldrich in the form of lyophilized powder and used without further purification. The purity for lysozyme and BSA is ≥ 98%, for myoglobin 95-100%, essentially salt-free. The properties and solvation dynamics of these proteins are well-studied by various experimental methods in the research community. The solutions are prepared using protein powder and distilled water with electrical resistivity of 13 MΩ·m. Solutions are prepared in nine different concentrations for each protein, namely

2.5, 5, 7.5, 10, 15, 20, 25, 50, 100 mg/ml. The most concentrated solution was prepared first; others were diluted from it using distilled water.

A Bruker liquid cell (A145) with TPX windows was utilized as a holder for the solutions. The thickness of the solutions was set to 100 μm by a polytetrafluoroethylene (PTFE) spacer. Since water is a strong absorber of sub-THz radiation, such thin samples allow measureable radiation through in the desired operating band while still being thick enough to provide sufficient beam-material interaction. Solutions are injected into the cell by a special pipette. The liquid cell is not opened while refilling. Instead it is firstly emptied with a syringe then washed with the next concentration, followed with injection of this concentration. This procedure is done to avoid any thickness deviation which is known to be a major source of uncertainty in absorption measurements [Sushko'13a, Duvillaret'99]. The temperature of the solutions is kept in the range 273-277K in an ice box. Before the measurement, solutions are kept outside the ice-box to allow them to equilibrate to room temperature.

Table 5-1: Protein properties that are exploited for interpretation of experimental data.

	Molecular weight, g/mole	Radius,[1] Å	Real protein dimensions, Å^3	Protein dipole moment,[2] Debye	Surface ratio[3]
Lysozyme, PDB: 2LYZ	14388	15.9	52x40x30	130	2.11
Myoglobin, PDB: 1MBN	17670	17.1	49x40x40	239	1.31
BSA, PDB: 4F5S	66500	27.1	95x75x60	844	1.51

[1]Computed from the partial specific volume and molar weight [Lee'83]

[2]Calculated from http://dipole.weizmann.ac.il

[3]The ratio of hydrophilic to hydrophobic solvent accessible area, estimated by Gromacs.

MD simulations were performed using the Gromacs package (version 4.5). Gromacs was utilized to accurately estimate the number of water molecules excluded by each protein and the number of water molecules in the protein interior. The simulation protocol is similar to that described in [Sushko'13b]. Proteins were solvated in a TIP3P water-box and allowed to equilibrate at 300 K before a production run of 100 ps. The required properties of protein and water were collected at the end of production run.

The complex transmission coefficients S_{21} of the samples were collected using a VNA-driven quasi-optical bench in the 220-325 GHz frequency domain. Two special frequency-extension heads, coupled to corrugated horns, were used as a receiver and transmitter. Two parabolic mirrors are used to focus emitted radiation on the sample. A further two parabolic mirrors direct the radiation to the receiving horn. Other details of the system can be found in [Yang'10]. The S_{21} amplitude, integrated over the whole investigated frequency band, is used as a measure of total transmittance or absorbance of the samples. The summed S_{21} amplitudes of solutions at each frequency point were normalized by those for water providing the relative changes. Only the relative absorbance of solutions with respect to buffer (water) is of interest in this study.

The measured S-parameters of the sample under test, S_{21} in particular, would unavoidably contain multiple FP-like reflections within the sample. CW radiation in the investigated band creates an interference pattern in the sample and is clearly detectable by VNA. However, no efforts have been made during data processing to de-embed [Hadjiloucas'13] multiple reflections from transmission coefficient for several reasons. First of all the water is a strong absorber at sub-THz frequencies therefore FP-echoes are highly attenuated even within the thin water sample of 100 μm. Also the TPX windows have relatively high surface roughness that enhances scattering and attenuates multiple reflections of radiation between window interfaces. Secondly, relatively long-range averaging has been applied to the readings that to some extent lessen the existing multiple reflections in the sample response. Finally the relative nature of the measurements (protein solution compared against water), compensates the effect of multiple reflections as well as effects of imperfect coupling, de-focusing of radiation, etc, since only the relative change in absorption and not the exact material properties of the sample is of interest in this study.

5.3 Absorption of solvated lysozyme, myoglobin and BSA proteins in 220-325 GHz band

The initial purpose of the experiments was to check whether all protein solutions consistently exhibit a decreased electromagnetic field absorption compared to a buffer (THz defect) over the considered sub-THz spectral domain. Fig. 1 shows the relative absorption of selected proteins at different concentrations. A nonlinear trend in absorption is immediately clear, especially at low concentration. Dashed lines represent the absorption of the solution excluding the absorption of water replaced by protein (two-component treatment, assuming protein absorption to be negligibly small). The

number of water molecules excluded by each protein has been estimated using the Gromacs MD package. The THz defect (i.e. the smaller absorption of solution compared to bulk water), was the expected scenario since protein molecules in the dry-state are known to be less absorbing than water. However, all three protein solutions exhibit higher absorption as compared to water and to two-component treatment (Fig. 5-1). Lysozyme is seen to have the largest additional absorption with respect to the simplified water-exclusion model. The absorption of BSA solution approaches the two-component model at high concentrations, however, it is not correct to conclude that the contribution from the protein molecules in particular is small, since the absorption of the hydration shell has to be considered. The proteins solutions studied exhibit an initial rise in absorption (THz excess) at low concentrations. The maximum absorbance of solutions occurs at 5 mg/ml, 7.5 mg/ml and 10 mg/ml for BSA, myoglobin and lysozyme respectively.

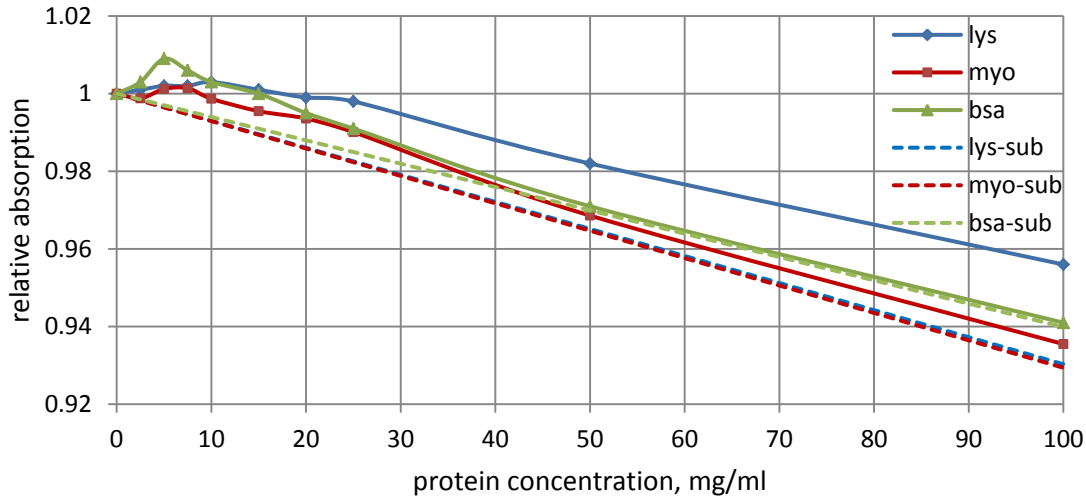


Figure 5-1. The relative concentration-resolved absorption of protein-water solutions at specific concentrations. Dashed lines represent the absorption of the solution excluding the absorption of water replaced by protein.

Only limited information regarding this feature in absorption by solutions is present in literature, therefore it needs further investigation. A similar observation has been made by Bye et al. [Bye'14], where the BSA-water solutions initially showed a rise or a plateau at low concentrations followed by a decrease in absorption. Previously, these peaks were related to the onset of the hydration shell overlap [Heyden'10a] and the hydration shell size was extracted accordingly. However in my study this seems inappropriate, since the depth of the hydration shells estimated at these concentrations was exceeding 50 Å. One of the possible explanations for THz excess at low

concentrations might be the process of distant interaction between protein dipole and collective the dipole of water molecules. Heyden et. al. [Heyden'12a] claimed that spatial-correlation between protein and water dipoles can extend out to 40-50 Å into the bulk water according to their computational study. If this holds, it might appear to be a reason for the absorption peak at specific concentrations. Initially at dilute concentrations, proteins build up this extended interaction with water comfortably. Then with inter-protein distance getting smaller, dipole interactions become less distant, resulting in absorption decrease. However this theory needs additional analysis in order to be proven. At present it remains challenging to unambiguously associate these peak positions with any physical processes. Instead we introduce a new approach towards extraction of the extent of the hydration water layer around protein, based on concentration-resolved absorption data of protein solutions.

Here the solution is treated by a three-component model as in [Heugen'06], so that the total absorption is governed by:

$$\alpha_s = \left(\alpha_p V_p + \alpha_w V_w + \alpha_h V_h\right)/V_s \tag{5.1}$$

where α and V are the absorption coefficient and volume of solution. Subscripts s, p w, and h are respectively, solution, protein, bulk water and hydration water. Note that contribution from dynamic protein-water hydrogen bonds to the total absorption are considered to be negligibly small.

The absorption by bulk water, and the solution, in the investigated frequency domain (220-325 GHz) is determined in our experimental setup. The volumes of components at each concentration are estimated using Gromacs. An initial assumption was made that the radius of the hydration shell is 10 Å. This initial guess does not influence in any way the calculation of the actual hydration shell radius suggested below. The relative absorption of hydration water is taken to constitute, on average, 90% of the bulk water absorption. According to MD simulations [Sushko'13b], only the first two water layers are highly perturbed by the protein, and the average response of whole hydration shell is dominated by more distant waters. Again the 10% value has only minor effect on the protein absorption in solution and no effect on the determination of the hydration shell size from measurements.

From (5.1) the absorption of protein α_p is determined. The concentration-dependent protein absorption in solution for the three proteins is plotted in Fig. 5-2. The data is presented as a ratio of protein absorption to water absorption *k*. For instance, given that water absorption at 270 GHz is 132 cm^{-1}, the absolute BSA absorption at the peak is 488 cm^{-1} (k=3.7). Errors were estimated as a combination of instrument noise

and the uncertainty of positioning of the liquid cell in the beam-line. Despite large errors at low concentrations, it is clear that he absorption of protein molecules experiences a sharp rise towards the dilution limit. These finding are in agreement with previously published results. For instance, Xu et al [Xu'06b] have determined the molar extinction of solvated BSA using a two-component model at 1.56 THz. They found that molar absorption reaches its steady-state at a protein concentration of around 4-5 % (approximately 40-50 mg/ml), in agreement with the findings here. Their data also shows a sharp rise in protein extinction at low concentrations; however the uncertainty is too high to make a firm conclusion. Zhang and Durbin did a similar analysis – determining the molar absorption coefficient of solvated myoglobin at different concentrations employing a three-component approach (accounting for bio-water) [Zhang'06]. The absorption curve also peaks at the dilution limit and reaches a plateau at a myoglobin concentration of approximately 7% at 0.35 THz.

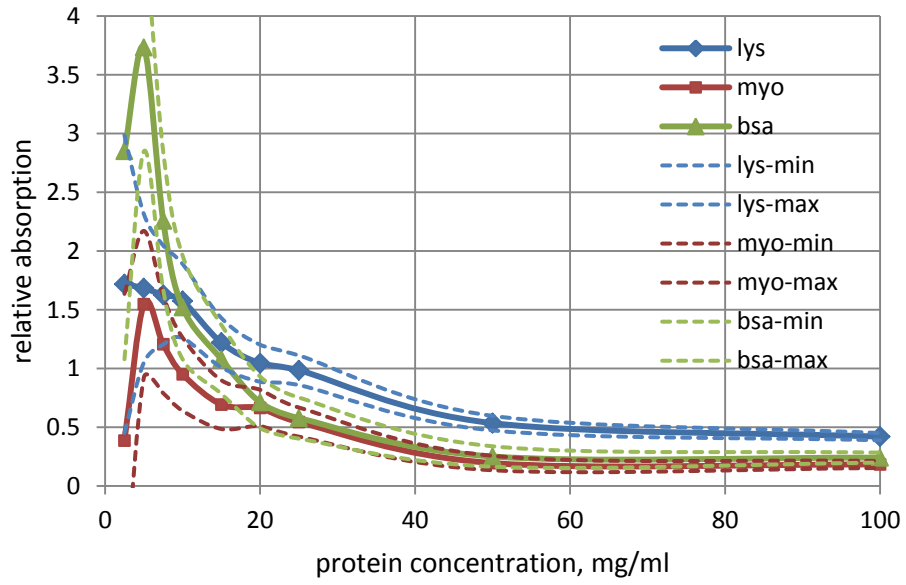


Figure 5-2. The absorbance of the solvated-protein molecules at different concentration with respective confidence bounds shown by dashed lines.

Proteins, and bio-molecules in general, in the dry-state (often in the form of lyophilized powders), are known to be far less absorbing compared to water in the THz and sub-THz frequency bands. The absorption coefficient of solvated protein (fig. 5-2) is shown to be comparable and even higher than that of bulk water at low protein concentrations. Absorption of electromagnetic radiation in the sub-THz band originates from the collective-dipole reorientation dynamics. Protein molecules in water solution have rotational relaxation time of the order of nanoseconds. This means that protein

dipole moment continuously lags behind the alternating electromagnetic field, producing a weak response to radiation in the 220-325 GHz domain (duty cycle 3-4 ps). However, the charged and hydrophilic side chains of amino acids interact more strongly with incident radiation in this frequency domain [Zhang'06]. At low concentrations, when inter-protein interaction is weak, and most of the water is bulk-like, these side-chains are especially dynamically active. This explains the underlying reasons for abnormally high absorption of proteins in solutions approaching the dilution limit. It is however challenging to estimate the proportion of side chains contributing to the total absorption in a solution, while seeing it has a noticeable effect.

5.4 Determination of hydration shell radius

At moderate protein concentrations of about 50 mg/ml, the average protein-to-protein distance is getting smaller, and most of remaining water is hydration water. At this concentration, protein molecules start stronger to interact with each other and the average dynamics of water diminishes due an increasing amount of retarded water in the solvation shells of proteins. This leads to hindrance of rotational and translational motions of charged and hydrophilic side chains of amino acids. As a result, beyond specific concentration the absorption coefficient levels off and remains constant for all three proteins (at least until 350 mg/ml for BSA, data not shown). Further, this concentration is referred as the 'critical' concentration. Interestingly, the steady absorption at critical concentration (0.42 – for lysozyme, 0.24 – for BSA and 0.18 for myoglobin), of the proteins considered here, is in quantitative agreement with the hydrophilic properties of proteins. The ratio of hydrophilic to hydrophobic solvent accessible areas (given in Table 5-1), results in the fact that, at any specific concentration, lysozyme molecules possesses the highest total hydrophilic area among the considered proteins, followed by BSA and myoglobin. This chemical property of the protein surface is found to be correlated to the protein absorption in solution. The 'critical' concentration marks the onset of overlapping of dynamic hydration shells of neighboring protein molecules in solution, and provides an alternative approach for the determination of hydration shell size.

Firstly, the average inter-protein distance has to be determined. The Gromacs MD package has been used to estimate the volume and the number of molecules in a cubic cell at each concentration. Two factors have taken into account of importance for hydration shell determination, which have not been considered before (to the best of my

knowledge). Firstly the proteins were not treated as idealized spherical particles. Instead the actual size of proteins (Table 5-1) has been taken into account. Next, the water molecules buried in the protein interior [Sushko'13b] were not considered in the calculation of the hydration shell size. Only the on-surface water molecules belong to the hydration shell in its conventional meaning. According to my approach, the final expression for estimation of hydration shell-size is governed by:

$$R_{shell} = 0.5\left(\left(N_{repl} + \frac{M_p}{M_w}\left(\frac{1}{c} - 1\right) - N_{int}\right)/\rho\right)^{1/3} - R_p, \tag{5.2}$$

where N_{repl} is the number of water molecules replaced by a single protein molecule; M_p, M_w are the molar weights of protein and water respectively; c is a concentration, expressed as a weight percentage of protein in solution; N_{int} is the number of water molecules that belong to a protein's interior structure as estimated by Gromacs (generally approximately equal to the protein residue number); ρ is a water density expressed in terms of the number of molecules per nm^3; R_p is a mean protein radius, calculated as an average of its x, y, z dimensions. The term in brackets is, effectively, the volume of a cubic box at a specific concentration. The cubic root of it respectively provides its side. The visual interpretation of the protein and its hydration shell is depicted in Fig. 5-3.

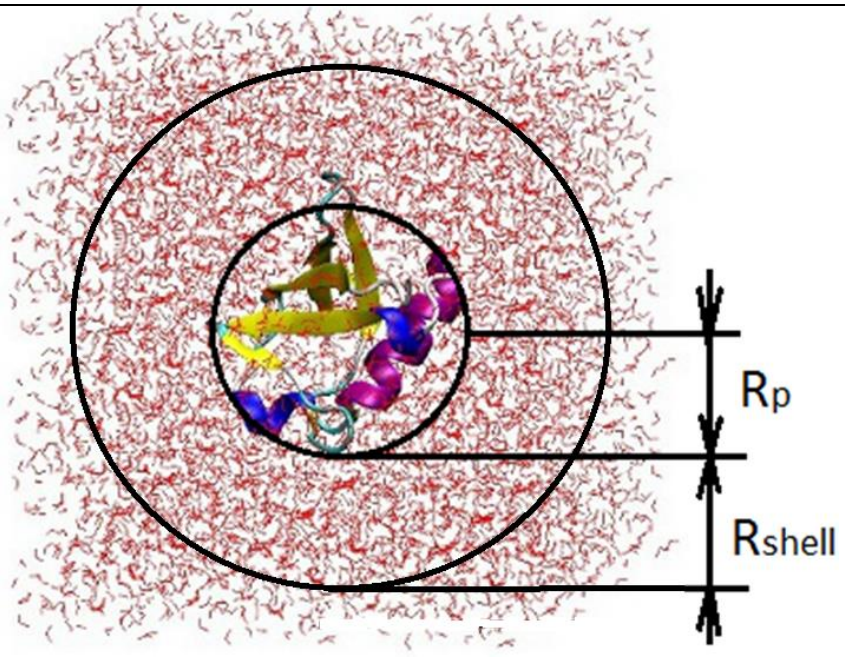


Figure 5-3. The graphical representation of the radii of hydration shells and protein.

The hydration shell radii calculated according to eqn. (5.2) are 16 Å, 19 Å and 25 Å for lysozyme, myoglobin and BSA respectively. Such a depth of hydration water

around proteins proves that protein-water interactions are complex and extend beyond first two to three layers of water molecules as estimated by MD simulations [Sushko'13b]. Importantly, the working algorithm does not depend on the absorption of hydration layers α_h, since it only effects the relative absorption of solvated protein and not the absorption curve shape. The figures obtained are in reasonable agreement with previous findings [Ebbinghaus'07, Ding'10]. For instance, Ding et. al. have estimated the hydration shell size for alanin-rich peptides to range from 11 to 17 Å, based on their THz spectrum. The radius of bio-water around λ-repressor and ubiquitin were calculated to reach 22 and 14 Å thickness respectively [Heyden'10a]. The reasons for an increasing hydration shell thickness with protein molar weight in my study, might originate from different dipole moments of proteins (Table 5-1). BSA, with its dipole moment of 844 Debye, can alter water dynamics further from the surface compared with lysozyme, having 6.5 times a smaller dipole moment. These agree with the observation that correlation between water collective-dipole and protein electric field, slowly build up only for distant water layers [Heyden'12a].

Applying the above analysis to the absorption data for BSA solutions from [Bye'14], leads to the interesting conclusion on the frequency-dependence of the hydration shell size. According to it, the hydration shell size is 12-14 Å (in close agreement with 15 Å as declared by authors [Bye'14]). The underlying reason for this difference in the hydration shell radius might be in the manner different frequencies *sense* the dipole perturbations in water caused by protein. It appears that electromagnetic radiation at 0.3 THz *feels* a more extended hydration shell compared to 1 THz radiation. The reason for this lies in the fact that a 1 THz field alternates faster than a 0.3 THz AC field, allowing less time for the collective-water dipole to respond. Concomitantly, the absorption response of hydration water is weaker at 1 THz. Note, the extracted thickness of the hydration layer as discussed above does not account for the protein-protein interactions strength, which might be different for the three proteins considered here.

5.5 Brief summary

The concentration-resolved study of lysozyme, myoglobin and BSA proteins has been performed quasi-optically in the 220-325 GHz waveguide band. The absorption of solutions was subdivided into contributions from protein molecules, water in hydration shell and bulk water. Absorption by protein molecules is found to increase rapidly, approaching the dilution-limit and to level off at about 50 mg/ml concentration. This

concentration was then used in alternative approach to extract the thickness of the hydration shell surrounding the proteins. Importantly, this approach does not depend on the absorption by the hydration shell. Hydration shell radii were determined to be 16 Å, 19 Å and 25 Å for lysozyme, myoglobin and BSA respectively, that, again, respectively corresponds to approximately 5, 6 and 8 water layers. This study suggests that protein-water interactions are more extended than was previously determined by MD simulations and originates from collective dipole moment interactions. Other experimental techniques used to study hydration water (e. g. NMR, X-ray diffraction, neutron scattering) also do not sense an extended thickness of hydration shell (as mentioned in chapter 4, page 64). A VNA-driven QO bench is able to monitor extended dipole interaction just as with THz-TDS, but with higher DR and SNR and being less complex in alignment and utilization.

Chapter 6. Conclusions and future work

In this thesis insights have been provided into data processing related to THz-TDS; molecular dynamics (MD) simulation tools for molecular-mechanical partial interpretation of experimental data that studied the solvation dynamics of a variety of bio-molecules whose selection spanned size and conformal complexity. The features of protein solvation dynamics (hydration shell radius, absorption of solvated proteins) are shown to correlate with its structural and chemical properties (surface hydrophobicity, dipole moment, size).

Original contributions in this thesis are:

- The introduction and quantifying of the errors in the materials parameters obtained from THz-TDS caused by utilization the approximated 'transfer' function in the extraction procedure. In addition all other sources of systematic and random of errors were rigorously treated;
- The clear distinction made between modulus-argument and real-imaginary approach of handling complex amplitude data in extraction of optical dispersion curves of solid and condensed soft-matter materials;
- The demonstration that the differences in solvation dynamics between considered proteins primarily originate from the constrained water molecules, buried in the interior of a protein; while dynamics of on-surface waters are approximately invariant for the four bio-molecules considered;
- According to MD simulations, the radius of the protein hydration shell is found to be 10 Å, and does not depend on the secondary structure of the protein (as judged by the vibrational density of states of the hydration layers);
- The protein absorptivity in solution exhibits a sharp rise towards the so-called, 'dilution limit'. It can be a few times greater than water absorption itself. The protein absorption curve also plateaus after a specific critical concentration, and this marks the onset of hydration shells overlap;
- Based on experimental absorption data in the 0.22-0.325 THz domain, the extracted hydration shell sizes are 16, 19 and 25 Å respectively for lysozyme, myoglobin and BSA proteins. It appears to be more extended compared to MD

simulation and other techniques, due to the apparent long-range interaction between dipoles belonging to the protein and those belonging to water.

However, many important contributions are still to be made in this area; namely, the broadband concentration- and frequency-resolved THz dielectric properties of many bio-molecules. This can unravel the features of interaction between water and solute. Also, bio-molecules in living cells are solvated, not only in water, but in various salts that were shown to noticeably alter the solvation dynamics. Significantly, the role of THz dynamics of chaotropic and kosmotropic salts and their structuring effects on solutes and water are often overlooked [Kaun'05]. A detailed investigation is, therefore, needed for the salt solutions of bio-molecules, especially their influence on the extent of the hydration shell and the structural stability of a bio-molecule.

Extra effort should be spent to further investigate over a wider frequency band the feature of THz excess of protein solutions, especially its dependence on the protein size and hydrophobicity. Previously, THz excess of protein solutions has been demonstrated only above 2 THz (groups of M. Havenith [Heyden'10a] and D. Cumming [Saha'12]), while our findings show solutions to exhibit an enhanced THz absorption with respect to bulk water in 0.22 – 0.325 THz frequency domain. Also, while a significant number of studies have been focused on the THz response of hydration water, the response of solvated proteins in the THz domain has not attracted as much attention. Only a few groups have attempted the analysis of the molar absorption of solvated proteins [Xu'06b, Zhang'06]. However as revealed their studies, as well as our analysis of protein activity, shows an abnormal behavior manifested in high THz absorption. This feature of protein dynamics has not yet been studied in detail.

The importance of protein-water interactions that are more extended (5-8 water layers) than was previously determined by MD simulations and originates from dipole moment interactions cannot be underestimated. Other experimental techniques used to study hydration water (e. g. NMR, X-ray diffraction, neutron scattering) do not sense an extended thickness of the hydration shell. The existence of a broader hydration shell than previously detected may have an important influence on molecular recognition events such as substrate binding and the formation of macromolecular complexes as the effective capture radius may be extended. There is also the possibility of preorganization of the interacting molecules at a longer distance that previously known.

Another perspective direction of research in THz area, only partly mention in the thesis, is the investigation of the interaction of THz radiation with bio-matter, especially at high power levels, where non-linear response occurs. THz waves become more wide-spread with consequence of living organism being more exposed to it. Only initial efforts have been made to determine the safe levels of exposure to THz radiation. Scientific community needs more evidence on whether THz waves can affect bio-molecules via vibrational (or micro-thermal) means or just thermally.

In regard to simulation, future work should include investigation of solvation dynamics of more complex proteins using the methodology presented in Chapter 4. While a quantum mechanical approach to the same problem is very computationally costly, it should benefit considerably, especially as regards interpretation of experimental results. A first step could combine quantum and molecular mechanical modelling, where some parts of a bio-molecule and water are treated by quantum mechanics and the rest by molecular mechanics. A separate attention should be paid to the methods of calculating the complex dielectric permittivity of bio-molecules and its solutions. Different approaches (for instance based on auto-correlation of total dipole moment, normal mode analysis, principal component analysis, vibrational density of states) need additional efforts to provide consistent results.

The work presented in this thesis covers many aspects of the highly interdisciplinary field that is THz spectrometry and unveils its complexity and versatility. In addition, many bio-inspired applications of THz radiation also require a validation against an already established orthogonal approach like circular dichroism or nuclear magnetic resonance spectroscopy. Applications of THz radiation will widen as it resolves hurdles and pitfalls peculiar to it, i.e. strong atmospheric absorption and lack of affordable and bright sources.

References

[Alexandrov'10] B. S. Alexandrov, V. Gelev, A.R. Bishop, A. Usheva, K.O. Rasmussen "DNA breathing dynamics in the presence of a terahertz field" *Physics Letters A* 374 (2010) 1214–1217

[Arikawa'08] Arikawa, T.; Nagai, M.; Tanaka, K. "Characterizing Hydration State in Solution Using Terahertz Time-domain Attenuated Total Reflection Spectroscopy" *Chem. Phys. Lett.* 2008, 457, 12-17

[Armstrong'12] Armstrong C. M. "The Truth about Terahertz" *IEEE Spectrum*, September 2012, pp. 36-41.

[Arora'12] Arun Arora, Trung Quan Luong, Matthias Kruger, Young Jun Kim, Chang-Hoon Nam, Andreas Manz and Martina Havenith "Terahertz-time domain spectroscopy for the detection of PCR amplified DNA in aqueous solution" *Analyst*, 2012, 137, 575

[Ashley'73] J. R. Ashley and F. M. Palka, "Transmission cavity and injection stabilization of an X-band transferred electron oscillator," in *IEEE MTT-S Int. Microwave Symp. Dig*., 1973, pp. 181-182

[Bandyopadhyay'06] Bandyopadhyay, S.; Chakraborty, S.; Bagchi, B. Exploration of the Secondary Structure Specific Differential Solvation Dynamics between the Native and Molten Globule States of the Protein HP-36. *J. Phys. Chem. B* 2006, 110, 20629-20634.

[Baxter'09] Baxter, J. B.; Schmuttenmaer, C. A. "Carrier Dynamics in Bulk ZnO. I. Intrinsic Conductivity Measured by Terahertz Time Domain Spectroscopy." *Phys. Rev. B* 2009, 80, 235205.

[Beard'02] M. C. Beard, G. M. Turner, and C. A. Schmuttenmaer, "Measuring intramolecular charge transfer via coherent generation of THz radiation" *J. Phys. Chem. A* 106, 878 (2002).

[Bertie'96] J. E. Bertie and Z. D. Lan, "Infrared intensities of liquids: The intensity of the OH stretching band of liquid water revisited, and the best current values of the optical constants of H_2O(1) at 25degrees C between 15 000 and 1 cm^{-1} ," *Appl. Spectrosc.*, vol. 50, no. 8, pp. 1047–1057, 1996.

[Blanchard'11] F. Blanchard, A. Doi, T. Tanaka, H. Hirori, H. Tanaka, Y. Kadoya, and K. Tanaka "Real-time terahertz near-field microscope" *Optics Express*, Vol. 19, Issue 9, pp. 8277-8284 (2011)

[Bock'10] Bock, J., et al., "Mammalian stem cells reporgramming in response to Terahertz Radiation" *PLoS Biol*, 2010. 5(12): p. e15806.

[Born'99] M. Born and E. Wolf, "Principles of Optics" *Cambridge University*, 1999.

[Bye'14] Jordan W. Bye, Stefano Meliga, Denis Ferachou, Gianfelice Cinque, J. Axel Zeitler,

and Robert J. Falconer “Analysis of the Hydration Water around Bovine Serum Albumin Using Terahertz Coherent Synchrotron Radiation” *J. Phys. Chem. A*, 2014, 118 (1), pp 83–88

[Bykhovski’10] Alexei Bykhovski, and Boris Gelmont, “The Influence of Environment on Terahertz Spectra of Biological Molecules”, *J Phys Chem B.* 114(38): 12349-57, 2010

[Cao’04] Z. W. Cao Y. Xue, L. Y. Han, B. Xie, H. Zhou, C. J. Zheng, H. H. Lin and Y. Z. Chen “MoViES: molecular vibrations evaluation server for analysis of fluctuational dynamics of proteins and nucleic acids”, *Nucleic Acids Res.* 2004 Vol. 32, issue 2, pp 679-685

[Castro-Camus’08] Castro-Camus, E.; Johnston, M. B. “Conformational Changes of Photoactive Yellow Protein Monitored by Terahertz Spectroscopy” *Chem. Phys. Lett.* 2008, 455, 289−292.

[Chakraborty’07] Chakraborty, S.; Sinha, S. K.; Bandyopadhyay, S. “Low-Frequency Vibrational Spectrum of Water in the Hydration Layer of Protein: A Molecular Dynamics Simulation Study” *J. Chem. Phys. B* 2007, 111, 13626−13631.

[Chaplin’06] Chaplin, M. “Opinion: Do we underestimate the importance of water in cell biology?” *Nat. Rev., Mol. Cell Bio.* 2006, 7, 861-866.

[Chen’07] J.-Y. Chen, J. R. Knab, S. Ye, Y. He, and A. G. Markelz, “Terahertz dielectric assay of solution phase protein binding,” *Appl. Phys. Lett.*, vol. 90, pp. 243901-1–243901-3, 2007.

[Chen’08] Zhanguo Chen, Jianxun Zhao, Yuhong Zhang, Gang Jia, Xiuhuan Liu, Ce Ren, Wenqing Wu, Jianbo Sun, Kun Cao, Shuang Wang and Bao Shi “Pockel’s effect and optical rectification in (111)-cut near-intrinsic silicon crystals” *Appl. Phys. Lett*. 92, 251111 (2008)

[Choi’12] Da-Hye Choi, Heyjin Son, Seonghoon Jung, Jaehun Park, Woong-Yang Park,Oh Sang Kwon, and Gun-Sik Park, “Dielectric relaxation change of water upon phase transition of a lipid bilayer probed by terahertz time domain spectroscopy” *J. Chem. Phys.* 137, 175101 (2012)

[Davies’11] Giles Davies, “Terahertz time-domain spectroscopy,” lecture in *ESoA THz course*, UPC, Spain, June, 2011

[Ding’10] Tao Ding, Ruoyu Li, J. Axel Zeitler, Thomas L. Huber, Lynn F. Gladden, Anton P. J. Middelberg and Robert J. Falconer “Terahertz and far infrared Spectroscopy of alanine-rich peptides having variable ellipticity” *Optics Express* 2010 Vol. 18, No. 26

[Ding’11] Ding, T.; Huber, T.; Middelberg, A. P. J.; Falconer, R. J. “Characterization of Low-Frequency Modes in Aqueous Peptides Using Far-Infrared Spectroscopy and Molecular Dynamics Simulation” *J. Phys. Chem. A* 2011, 115, 11559−11565.

[Dorney’01] T. D. Dorney, R. G. Baraniuk, and D. M. Mittleman, "Material parameter estimation with terahertz time-domain spectroscopy," *J. Opt. Soc. Am. A,* vol. 18, pp. 1562-1571, 2001

[Duvillaret'96] L. Duvillaret, F. Garet, and J. L. Coutaz, "A reliable method for extraction of material parameters in terahertz time-domain spectroscopy," *IEEE Journal of Selected Topics in Quantum Electronics,* vol. 2, pp. 739-746, 1996.

[Duvillaret'99] L. Duvillaret, F. Garet, and J.-L. Coutaz, "Highly Precise Determination of Optical Constants and Sample Thickness in Terahertz Time-Domain Spectroscopy," *Appl. Opt.,* vol. 38, pp. 409-415, 1999.

[Duvillaret'00] L. Duvillaret, F. Garet, and J.-L. Coutaz, "Influence of noise on the characterization of materials by terahertz time-domain spectroscopy," *J. Opt. Soc. Am. B,* vol. 17, pp. 452-461, 2000.

[Ebbinghaus'07] Simon Ebbinghaus, Seung Joong Kim, Matthias Heyden, Xin Yu, Udo Heugen, Martin Gruebele, David M. Leitner, and Martina Havenith "An extended dynamical hydration shell around proteins" *PNAS,* 2007, vol. 104, no. 52, 20749–20752

[Ebbinghaus'08] Simon Ebbinghaus, Seung Joong Kim, Matthias Heyden, Xin Yu, Martin Gruebele, David M. Leitner, and Martina Havenith "Protein Sequence- and pH-Dependent Hydration Probed by Terahertz Spectroscopy" *J. Am. Chem. Soc.* 2008, 130, 2374-2375

[Ebbinghaus'10] Simon Ebbinghaus, Konrad Meister, Benjamin Born, Arthur L. DeVries, Martin Gruebele and Martina Havenith "Antifreeze Glycoprotein Activity Correlates with Long-Range Protein−Water Dynamics" *J. Am. Chem. Soc.*, 2010, 132 (35), pp 12210–12211

[Exeter'89] M. van Exeter, C. Fattinger and D. Grischkowsky, "Terahertz time-domain spectroscopy of water vapor," *Opt. Lett.*14, 1128 (1989).

[Exeter'90a] Van Exeter M. and Grischkowsky D. "Carrier dynamics of electrons and holes in moderately doped silicon" *Physical Review B*, 1990 41(17), pp. 12140–12149.

[Exeter'90b] Van Exeter M. and Grischkowsky D. "Optical and electronic properties of doped silicon from 0.1 to 2 THz" *Applied Physics Letters*, 1990, 56(17), pp. 1694–1696.

[Falconer'12] Robert J. Falconer, Andrea G. Markelz "Terahertz Spectroscopic Analysis of Peptides and Proteins" *J Infrared Milli Terahz Waves* (2012) 33:973–988

[Federici'05] J. F. Federici, B. Schulkin, F. Huang, D. Gary, R. Barat, F. Oliveira, and D. Zimdars, "THz imaging and sensing for security applications—explosives, weapons and drugs", *Semicond. Sci. Technol.* 20, 266–280, (2005).

[Fleming'74] J. W. Fleming, "High resolution submillimeter-wave Fourier-transform spectrometry of gases," *IEEE Trans. Microwave Theory Tech*., vol. MTT-22, pp. 1023–1025, 1974.

[Franz'08] Morten Franz, Bernd M. Fischer, and Markus Walther " The Christiansen effect in terahertz time-domain spectra of coarse-grained powders" *Appl. Phys. Lett.* 92, 021107, 2008

[Frohlich'75] Frohlich, H., "The extraordinary dielectric properties of biological materials and the action of enzymes" *Proc Natl Acad Sci USA*, 1975. 72(11): p. 4211–5.

[Frolich'08] Frolich A., Gabel F., Jasnin M., Lehnert U., Oesterhelt D. "From shell to cell: neutron scattering studies of biological water dynamics and coupling to activity" *Faraday Discussions*, 141, 2008, 117-130.

[Gekle'12] Stephan Gekle, Roland R. Netz "Anisotropy in the dielectric spectrum of hydration water and its relation to water dynamics" *Journal of Chemical Physics* 137, 104704 (2012)

[George'08] Paul A. George, Wallace Hui, Farhan Rana, Benjamin G. Hawkins, A. Ezekiel Smith, Brian J. Kirby "Microfluidic devices for terahertz spectroscopy of bio-molecules" *Optics Express* 2008, vol. 16, No. 3, 1577-1582.

[Glancy'10] P. Glancy, W. P. Beyermann "Dielectric properties of fully hydrated nucleotides in the terahertz frequency range" J. Chem. Phys. 132, 245102, 2010.

[Gretel'09] Gretel M. P., Robyn Flook, Brian W.-H., and Derek Abbott, "Terahertz Spectroscopy of Misfolded Proteins in Bio-Tissue", *IRMMW-THz* conference proceedings, 2009

[Hadjiloucas'99] S. Hadjiloucas, L. S. Karatzas and J.W. Bowen. "Measurements of Leaf Water Content Using Terahertz Radiation," *IEEE Transactions on Microwave Theory and Techniques MTT*, 47, 142-149 (1999)

[Hadjiloucas'13] S. Hadjiloucas, G. C. Walker and J. W. Bowen. "A 1-Port De-embedding Technique for the Quasi-optical Characterization of Integrated Components," *IEEE Sensors Journal* 13(1), pp. 111-123, 2013.

[He'08] Yunfen He, Pei I. Ku, J. R. Knab, J.Y. Chen, and A. G. Markelz "Protein Dynamical Transition Does Not Require Protein Structure" *PRL* 101, 178103 (2008)

[He'10] Yunfen He, J.-Y. Chen, J. R. Knab, Wenjun Zheng and A. G. Markelz, "Why is THz Sensitive to Protein Functional States? Oxidation State of Cytochrome C" *TST Conf. Proc.*, pp. 149-162, 2010

[He'11] Yunfen He, J.-Y. Chen, J. R. Knab, Wenjun Zheng, and A. G. Markelz "Evidence of Protein Collective Motions on the Picosecond Timescale" *Biophysical Journal* Volume 100, 2011 1058–1065

[Heugen'06] U. Heugen, G. Schwaab, E. Brundermann, M. Heyden, X. Yu, D. M. Leitner, and M. Havenith, "Solute-induced retardation of water dynamics probed directly by terahertz spectroscopy" *PNAS* August 15, 2006, vol. 103, no. 33, 12301–12306

[Heyden'08] M. Heyden, E. Bründermann, U. Heugen, G. Niehues, D. M. Leitner, and M. Havenith "Long-Range Influence of Carbohydrates on the Solvation Dynamics of WatersAnswers from Terahertz Absorption Measurements and Molecular Modeling Simulations" *J. Am. Chem. Soc.* 2008, *130*, 5773–5779

[Heyden'10a] Heyden, M.; Havenith, M. "Combining THz Spectroscopy and MD Simulations to Study Protein-Hydration Coupling" *Methods* 2010, 52, 74−83.

[Heyden'10b] Heyden, M.; Sun, J.; Funkner, S.; Mathias, G.; Forbert, H.; Havenith, M.; Marx, D. "Dissecting the THz Spectrum of Liquid Water from First Principles via Correlations in Time and Space" *Proc. Natl. Acad. Sci. U.S.A.* 2010, 107, 12068−12073.

[Heyden'12a] Matthias Heyden, Douglas J. Tobias, and Dmitry V. Matyushov "Terahertz absorption of dilute aqueous solutions" *J. Chem. Phys*. 137, 235103 (2012).

[Heyden'12b] Heyden, M.; Sun, J.; Funkner, S.; Mathias, G.; Forbert, H.; Havenith, M.; Marx, D. "Understanding the Origins of Dipolar Couplings and Correlated Motion in the Vibrational Spectrum of Water" *J. Phys. Chem. Lett.* 2012, 3, 2135-2140.

[Jeon'97] T. I. Jeon and D. Grischkowsky, "Nature of conduction in doped silicon" *Physical Review Letters*, 1997, 78(6), pp. 1106–1109.

[Jeon'98] T. I. Jeon and D. Grischkowsky, "Characterization of optically dense, doped semiconductors by reflection THz time domain spectroscopy" *Appl. Phys. Lett*. 72, 3032 (1998).

[Jepsen'05] P. U. Jepsen and B. Fischer, "Dynamic range in terahertz time-domain transmission and reflection spectroscopy," *Opt. Lett.*, vol. 30, pp. 29-31, 2005.

[Jepsen'11] Peter Uhd Jepsen, David G. Cooke, and Martin Koch "Terahertz spectroscopy and imaging – Modern techniques and applications" *Laser Photonics Rev.* 5, No. 1, 124–166 (2011)

[Jin'10] Biaobin Jin, Cunlin Zhang, Peiheng Wu, Shenggang Liu, "Recent Progress of Terahertz Spectroscopy on Medicine and Biology in China" *Terahertz Science and Technology*, Vol.3, No.4, 2010, 192-200.

[Jin'06] Yun-Sik Jin, Geun-Ju Kim and Seok-Gy Jeon "Terahertz Dielectric Properties of Polymers" *Journal of the Korean Physical Society*, Vol. 49, No. 2, August 2006, pp. 513-517

[Kambara'10] Ohki Kambara and Keisuke Tominaga "Structural fluctuation of proteins revealed by terahertz time-domain spectroscopy" *Spectroscopy* 24 (2010) 149-152.

[Kaun'05] N. Kaun, J. R. Baena, D. Newnham, and B. Lendl "Terahertz pulsed spectroscopy as a new tool for measuring the structuring effect of solutes on water," Applied Spectroscopy, vol. 59, no. 4, 2005, pp. 505-510.

[Kaushik'12a] Mayank Kaushik, Brian W.-H. Ng, Bernd M. Fischer, and Derek Abbott "Terahertz fingerprinting in presence of quasi-ballistic scattering", *Appl. Phys. Lett*. 101, 061108 (2012)

[Kaushik'12b] Mayank Kaushik, Brian W.-H. Ng, , Bernd M. Fischer,and Derek Abbott, "Reduction of Scattering Effects in THz-TDS Signals" *IEEE Photonics Technology Letters*, Vol. 24, No. 2, 2012

[Kawaguchi'10] Shintaro Kawaguchi, Ohki Kambara, Mikihiro Shibata, Hideki Kandori and Keisuke Tominaga "Low-frequency dynamics of bacteriorhodopsin studied by terahertz time-domain spectroscopy" *Phys. Chem. Chem. Phys*., 2010, 12, 10255–10262

[Kemp'06] Michael C. Kemp, "Millimetre Wave and Terahertz Technology for the Detection of Concealed Threats – A Review", *Proc. of SPIE* Vol. 6402, 64020D, (2006)

[Kersting'08] R. Kersting, Federico F. Buersgens, G. Acuna, G. C. Cho "Terahertz near-field microscopy", *Adv. In Solid State Phys*. 47, 203-222 (2008).

[King'11] M. D. King, W. D. Buchanan, T. M. Korter, "Understanding the Terahertz Spectra of Crystalline Pharmaceuticals: Terahertz Spectroscopy and Solid-State Density

Functional Theory Study of (S)-(+)-Ibuprofen and (RS)-Ibuprofen", *J. Pharm. Sci*., 100(3), 1116-1129, 2011

[Knab'06] Joseph Knab, Jing-Yin Chen, and Andrea Markelz "Hydration Dependence of Conformational Dielectric Relaxation of Lysozyme" *Biophysical Journal* Vol. 90, 2006 2576–2581

[Knab'07] Joseph R. Knab, Jing-Yin Chen, Yunfen He, and Andrea G. Markelz "Terahertz Measurements of Protein Relaxational dynamics" *Proceedings of the IEEE*, vol. 95, No. 8, 2007

[Krüger'11] M. Krüger, S. Funkner, E. Bründermann, and M. Havenith, "Uncertainty and Ambiguity in Terahertz Parameter Extraction and Data Analysis," *Journal of Infrared, Millimeter and Terahertz Waves*, vol. 32, pp. 699-715, 2011.

[Kubelka'04] Kubelka, J.; Hofrichter, J.; Eaton, W. A. "The Protein Folding 'Speed Limit'". *Curr. Opin. Struct. Biol.* 2004, 14, 76-88

[Laurette'10] S. Laurette, A. Treizebre, F. Affouard and B. Bocquet, "Subterahertz characterization of ethanol hydration layers by microfluidic system" *Appl. Phys. Lett*., 2010, 97, 111904.

[Laurette'12] Simon Laurette, Anthony Treizebre, Adil Elagli, Basak Hatirnaz, Renato Froidevaux, Frederic Affouard, Ludovic Duponchelb and Bertrand Bocquet "Highly sensitive terahertz spectroscopy in microsystem" *RSC Advances*, 2012, 2, 10064–10071

[Lee'83] B. Lee "Calculation of volume fluctuation for globular protein models" *Proc. Natl. Acad. Sci. USA* Vol. 80, pp. 622-626, 1983

[Leitner'08] David M Leitner, Martin Gruebele and M. Havenith, "Solvation dynamics of bio-molecules: modeling and terahertz experiments", *HFSP Journal*, Vol. 2, no. 6, 314-323, 2008.

[Loewenstein'96] E. V. Loewenstein, "The history and current status of Fourier Transform spectroscopy," *Appl. Opt*. 5, 845 (1966).

[Liu'04] Liu, P.; Harder, E.; Berne B. J. "On the Calculation of Diffusion Coefficients in Confined Fluids and Interfaces with an Application to the Liquid-Vapor Interface of Water" *J. Phys. Chem. B* 2004, 108, 6595-6602.

[Lindorff-Larsen'11] Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Shaw, D. E. "How Fast-folding Proteins Fold" *Science* 2011, 334, 517-520

[Marchi'02] Marchi, M.; Sterpone, F.; Ceccarelli, M. "Water Rotational Relaxation and Diffusion in Hydrated Lysozyme" *J. Am. Chem. Soc*.2002, 124, 6787-6791.

[Markelz'00] A.G. Markelz, A. Roitberg, E.J. Heilweil "Pulsed terahertz spectroscopy of DNA, bovine serum albumin and collagen between 0.1 and 2.0 THz" *Chemical Physics Letters* 320 (2000) 42–48

[Markelz'02] Andrea Markelz, Scott Whitmire, Jay Hillebrecht and Robert Birge "THz time domain spectroscopy of biomolecular conformational modes" *Phys. Med. Biol*. 47 (2002) 3797–3805

[Markelz'07] Andrea G. Markelz, Joseph R. Knab, Jing Yin Chen, Yunfen He "Protein dynamical transition in terahertz dielectric response" *Chemical Physics Letters* 442 (2007) 413–417

[Markelz'08] Andrea G. Markelz "Terahertz Dielectric Sensitivity to Biomolecular Structure and Function" *IEEE Journal of Selected Topics in Quantum Electronics*, Vol. 14, No. 1, 2008

[Matvejev'12] Matvejev, V.; Zizi, M.; Stiens, J. J. "Hydration Shell Parameters of Aqueous Alcohols: THz Excess Absorption and Packing Density" *J. Phys. Chem. B*2012, 116, 14071-14077.

[Mattea'08] Mattea, C.; Qvist, J.; Halle, B. "Dynamics at the Protein-Water Interface from ^{17}O Spin Relaxation in Deeply Supercooled Solutions" *Biophys. J.* 2008, 95, 2951–2963

[Mickan'04] S. Mickan, Jingzhou Xu, J. Munch, X.-C. Zhang, and D. Abbott, "The limit of spectral resolution in THz time-domain spectroscopy", *Proc. of SPIE*, Vol. 5277, 2004.

[Naftaly'05] M. Naftaly, A. P. Foulds, R. E. Miles, and A. G. Davies "Terahertz Transmission Spectroscopy of Nonpolar Materials and Relationship with Composition and Properties" *International Journal of Infrared and Millimeter Waves*, Vol. 26, No. 1, 2005

[Naftaly'07a] Mira Naftaly and Robert E. Miles, "Terahertz Time-Domain Spectroscopy for Material Characterization" *Proceedings of the IEEE*, Vol. 95, No. 8, 2007

[Naftaly'07b] M. Naftaly and R. E. Miles, "A method for removing etalon oscillations from THz time-domain spectra," *Optics Communications*, vol. 280, pp. 291-295, 2007.

[Naftaly'09] M. Naftaly and R. Dudley, "Methodologies for determining the dynamic ranges and signal-to-noise ratios of terahertz time-domain spectrometers," *Optics Letters*, vol. 34, no. 8 pp. 1213-1215, 2009.

[Nagai'06] M. Nagai, H. Yada, T. Arikawa, and K. Tanaka, "Terahertz time-domain attenuated total reflection spectroscopy in water and biological solution," *Int. J. Infrared Millim. Waves* 27, 505-515 (2006).

[Niehues'11] Gudrun Niehues, Matthias Heyden, Diedrich A. Schmidt and Martina Havenith "Exploring hydrophobicity by THz absorption spectroscopy of solvated amino acids" *Faraday Discuss*., 2011, 150, 193–207

[Pal'13] Pal, S.; Bandyopadhyay, S. "Effects of Protein Conformational Flexibilities and Electrostatic Interactions on the Low-Frequency Vibrational Spectrum of Hydration Water" *J. Phys. Chem. B* 2013, 117, 5848−5856.

[Pertsemlidis'96] A. Pertsemlidis, A. M. Saxena, A. K. Soper, T. Head-Gordon, and R. M. Glaeser, "Direct evidence for modified solvent structure within the hydration shell of a hydrophobic amino acid" *Proc. Natl. Acad. Sci. U.S.A.* 1996, Vol. 93, No. 20 10769–10774.

[Pronk'13] Pronk, S.; Pall, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; van der Spoel, D.; Hess, B.; Lindahl, E. "GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit" *Bioinformatics* 2013, 29, 845–854.

[Pupeza'07] I. Pupeza, R. Wilk, and M. Koch, "Highly accurate optical material parameter determination with THz time-domain spectroscopy," *Opt. Express*, vol. 15, pp. 4335-4350, 2007.

[Qvist'09] Qvist, J.; Persson, E.; Mattea, C.; Halle, B. "Time Scales of Water Dynamics at Biological Interfaces: Peptides, Proteins and Cells" *Faraday Discuss.*2009, 141, 131–144.

[Rocchi'98] Rocchi, C.; Bizzarri, A. R.; Cannistraro, S. "Water Dynamics Anomalies Evidenced by Molecular-Dynamics Simulations at the Solvent-Protein Interface" *Phys. Rev. E* 1998, 57, 3315-3325.

[Saha'12] S. C. Saha, James P. Grant, Yong Ma, Ata Khalid, Feng Hong, and David R. S. Cumming "Application of terahertz spectroscopy to the characterization of biological samples using birefringence silicon grating," *Journal of Biomedical Optics* 17(6), 067006 (2012)

[Scheller'09a] Scheller M., Jancen C. and Koch M. "Analyzing sub-100-µm samples with transmission terahertz time domain spectroscopy" 2009, *Opt. Comm.* 282, 1304-06.

[Scheller'09b] Scheller M. and Koch M. "Fast and Accurate Thickness Determination of Unknown Materials using Terahertz Time Domain Spectroscopy" 2009 *J. Infrared Mili Terahz Waves* Vol. 30, pp. 30762-769.

[Sengupta'08] Sengupta, N.; Jaud, S.; Tobias, D. J. "Hydration Dynamics in a Partially Denatured Ensemble of the Globular Protein α-Lactalbumin Investigated with Molecular Dynamics Simulations" *Biophys. J.* 2008, 95, 5257-5267.

[Shan'04] Jie Shan and Tony F. Heinz "Terahertz Radiation from Semiconductors" *Topics Appl. Phys.* 92, 1–59 (2004)

[Shen'08] Y. C. Shen, P. F. Taday, and M. Pepper "Elimination of scattering effects in spectral measurement of granulated materials using terahertz pulsed spectroscopy" *Appl. Phys. Lett*. 92, 051103 (2008).

[Siegel'02] Peter H. Siegel, "Terahertz Technology," *IEEE Transanctions on MTT*, Vol. 50, No. 3, pp. 910 - 928, Mar 2002.

[Siegel'04] Peter H. Siegel, "Terahertz Technology in Biology and Medicine," *IEEE Transanctions on MTT*, Vol. 52, No. 10, pp. 2438 - 2447, Oct 2004.

[Sinha'08] Sinha, S. K.; Chakraborty, S.; Bandyopadhyay, S. "Thickness of the Hydration Layer of a Protein from Molecular Dynamics Simulation" *J. Phys. Chem.* 2008, 112, 8203-8209.

[Sinha'11] Sinha, S. K.; Bandyopadhyay, S. "Differential Flexibility of the Secondary Structures of Lysozyme and the Structure and Ordering of Surrounding Water Molecules" *J. Chem. Phys.* 2011, 134, 115101.

[Sinha'12a] Sinha, S. K.; Bandyopadhyay, S. "Polar Solvation Dynamics of Lysozyme from Molecular Dynamics Studies" *J. Chem. Phys.* 2012, 136, 185102.

[Sinha'12b] Sinha, S. K.; Bandyopadhyay, S. "Local heterogeneous dynamics of water around lysozyme: a computer simulation study" *Phys. Chem. Chem. Phys.*2012, 14, 899–913.

[Sushko'13a] Oleksandr Sushko, Kastriot Shala, Rostyslav Dubrovka, and Robert Donnan "Revised metrology for enhanced accuracy in complex optical constant determination by THz-time-domain spectrometry" *J. Opt. Soc. Am. A*, Vol. 30, No. 5, 979-986, 2013

[Sushko'13b] Oleksandr Sushko, Rostyslav Dubrovka, and Robert S. Donnan "Terahertz Spectral Domain Computational Analysis of Hydration Shell of Proteins with Increasingly Complex Tertiary Structure" *J. Phys. Chem. B* 2013, 117, 16486−16492

[Svergun'98] D. I. Svergun, S. Richard, M. H. J. Koch, Z. Sayers, S. Kuprin, G. Zaccai, "Protein hydration in solution: Experimental observation by x-ray and neutron scattering" *Proc. Natl. Acad. Sci. USA* Vol. 95, pp. 2267–2272, 1998

[Swanson'11] Eric S. Swanson "Modeling DNA response to terahertz radiation" *Phys. Rev. E* 83, 040901(R) (2011)

[Tonochi'07] Masayoshi Tonochi, "Cutting-edge Terahertz Technology," *Nature*, Vol. 1, pp. 97 - 105, Feb, 2007.

[Torben'10] Torben T. L. Kristensen, Withawat Withayachumnankul, Peter Uhd Jepsen and Derek Abbott "Modeling terahertz heating effects on water" *Optics Express* 2010, Vol. 18, No. 5, 4727-4739

[Ueno'08] Yuko Ueno and Katsuhiro Ajito, "Analytical Terahertz spectroscopy," *Analytical Sciences*, Vol. 24, pp 185-192, 2008.

[Vinh'11] Vinh, N. Q.; Allen, S. J.; Plaxco, K. W. "Dielectric Spectroscopy of Proteins as a Quantitative Experimental Test of Computational Models of Their Low-frequency Harmonic Motions" *J. Am. Chem. Soc.* 2011, 133, 8942-8947

[Whitmire'03] S. E. Whitmire, D. Wolpert, A. G. Markelz, J. R. Hillebrecht, J. Galan, and R. R. Birge "Protein Flexibility and Conformational State: A Comparison of Collective Vibrational Modes of Wild-Type and D96N Bacteriorhodopsin" *Biophysical Journal* Volume 85, 2003, 1269–1277

[Wilmink'11] Gerald J. Wilmink, Jessica E. Grundt "Invited Review Article: Current State of Research on Biological Effects of Terahertz Radiation", *J Infrared MilliTerahz Waves* (2011) 32:1074–1122

[Woodward'03] R. M. Woodward, V. P. Wallace, D. D. Arnone, E. H. Linfield, and M. Pepper, "Terahertz pulsed imaging of skin cancer in the time and frequency domain," *J. Biol. Phys.*, vol. 29, no. 2/3, pp. 257–259, 2003.

[Wilk'08] R. Wilk, I. Pupeza, R. Cernat, and M. Koch, "Highly accurate thz time-domain spectroscopy of multilayer structures", *IEEE Journal of Selected Topics in Quantum Electronics*, *14*(2), 392-398, 2008.

[Winnewisser'97] C. Winnewisser, P. Uhd Jepsen, M. Schall, V. Schyja and H. Helm "Electro-optic detection of THz radiation in LiTaO3, LiNbO3 and ZnTe" *Appl. Phys. Lett.* 70, 3069 (1997);

[Withayachumnankul'05] W. Withayachumnankul, B. Ferguson, T. Rainsford, S. P. Mickan, and D. Abbott, "Simple material parameter estimation via terahertz time-domain spectroscopy," *Electronics Letters*, vol. 41, pp. 800-801, 2005.

[Withayachumnankul'07] W. Withayachumnankul, H. Lin, S. P. Mickan, B. M. Fischer, D. Abbott, "Analysis of measurement uncertainty in THz-TDS," *Proc. of SPIE* vol. 6593, 2007.

[Withayachumnankul'08a] W. Withayachumnankul, B. M. Fischer, H. Lin, and D. Abbott, "Uncertainty in terahertz time-domain spectroscopy measurement," *J. Opt. Soc. Am. B,* vol. 25, pp. 1059-1072, 2008.

[Withayachumnankul'08b] Withawat Withayachumnankul, Bernd M. Fischer and Derek Abbott, "Numerical removal of water vapour effects from terahertz time-domain spectroscopy measurements" *Proc. R. Soc. A* (2008) 464, 2435–2456

[Xie'12] Xie W X, Li J, Pei J H., "THz-TDS Signal Analysis and Substance Identification via the Conformal Split" *Sci. China Inf. Sci.* vol. 55, pp. 49–63, 2012.

[Xu'06a] Jing Xu, Kevin W. Plaxco, and S. James Allen, "Collective Dynamics of Lysozyme in Water: Terahertz Absorption Spectroscopy and Comparison with Theory", *J. Phys. Chem. B*, 110 (47), pp 24255–24259, 2006.

[Xu'06b] Jing Xu, Kevin W. Plaxco, and S. James Allen "Probing the collective vibrational dynamics of a protein in liquid water by terahertz absorption spectroscopy" *Protein Science* (2006), 15:1175–1181

[Xu'12] Xu, Y.; Gnanasekaran, R.; Leitner, D. M. "Analysis of Water and Hydrogen Bond Dynamics at the Surface of an Antifreeze Protein" *J. At. Mol. Opt. Phys.* 2012, 2012, 125071.

[Yada'08] Hiroyuki Yada, Masaya Nagai, Koichiro Tanaka "Origin of the fast relaxation component of water and heavy water revealed by terahertz time-domain attenuated total reflection spectroscopy" *Chemical Physics Letters* 464 (2008) 166–170

[Yang'10] Bin Yang, Richard J. Wylde, Derek. H. Martin, Philippe Goy, Robert S. Donnan, and Sylvain Caroopen "Determination of the Gyrotropic Characteristics of Hexaferrite Ceramics From 75 to 600 GHz" *IEEE Transactions on MTT*, Vol. 58, No. 12, 2010

[Yamaguchi'05] M. Yamaguchi, F. Miyamaru, K Yamamoto, M. Tani, M. Hangyo, "Terahertz absorption spectra of L-, D-, and DL-alanine and their application to determination of enantiometric composition" *Appl. Phys. Lett*. 2005, 86, 053903

[Yamamoto'97] Naoki Yamamoto, Ohki Kambara, Kohji Yamamoto, Atsuo Tamura, Shinji Saito and Keisuke Tominaga "Temperature and hydration dependence of low-frequency spectra of poly-L-glutamic acid with different secondary structures studied by terahertz time-domain spectroscopy" *Soft Matter*, 2012, 8, 1997

[Yamamoto'12] Naoki Yamamoto, Ohki Kambara, Kohji Yamamoto, Atsuo Tamura, Shinji Saito and Keisuke Tominaga "Temperature and hydration dependence of low-frequency spectra of poly-L-glutamic acid with different secondary structures studied by terahertz time-domain spectroscopy" *Soft Matter*, 2012, 8, 1997-2006

[Zeitler'07] J. Axel Zeitler, Philip F. Taday, David A. Newnham, M. Pepper, Keith C. Gordon and T.Rades "Terahertz pulsed spectroscopy and imaging in the pharmaceutical setting – a review" *JPP* 2007, 59: 209–223.

[Zhang'06] Chenfeng Zhang and Stephen M. Durbin "Hydration-Induced Far-Infrared Absorption Increase in Myoglobin" *J. Phys. Chem. B* 2006, *110,* 23607-23613

[Zhang'09] Zhang X. C., J. Xu, "Introduction to THz wave photonics" *Springer*, 2009