TB, RMK, UK, PMB/349644, 23/07/2010

IOP PUBLISHING PHYSICS IN MEDICINE AND BIOLOGY

Phys. Med. Biol. **55** (2010) 1–14 UNCORRECTED PROOF

Accuracy and resolution of THz reflection spectroscopy for medical imaging

Caroline B Reid[1], Emma Pickwell-MacPherson[2], Jan G Laufer[1], Adam P Gibson[1], Jeremy C Hebden[1] and Vincent P Wallace[3]

[1] Department of Medical Physics and Bioengineering, University College London, London WC1E 6BT, UK
[2] Electronic and Computer Engineering Department, The Hong Kong University of Science and Technology, Clearwater Bay, Hong Kong
[3] School of Physics, The University of Western Australia, Crawley 6009, Australia

E-mail: c.reid@medphys.ucl.ac.uk

Received 2 March 2010, in final form 3 March 2010
Published DD MMM 2010
Online at stacks.iop.org/PMB/55/1

Abstract

The use of THz radiation as a potential tool for medical imaging is of increasing interest. In this paper three methods of analysis of THz spectroscopic information for diagnosis of tissue pathologies at THz frequencies are presented. The frequency-dependent absorption coefficients, refractive indices and Debye relaxation times of pure water and pure lipids were measured and used as prior knowledge in the different theoretical methods for the determination of concentration. Three concentration analysis methods were investigated: (a) linear spectral decomposition, (b) spectrally averaged dielectric coefficient method and (c) the Debye relaxation coefficient method. These methods were validated on water and lipid emulsions by determining the concentrations of phantom chromophores and comparing to the known composition. The accuracy and resolution of each method were determined to assess the potential of each method as a tool for medical diagnosis at THz frequencies.

(Some figures in this article are in colour only in the electronic version) Q1

1. Introduction

The terahertz (THz = 10^{12} Hz) range lies between the microwave and infrared regions of the electromagnetic spectrum, with a wavelength and frequency range, typically, defined as 0.3–3 mm and 0.1–10 THz, respectively (Siegel 2004). THz radiation is non-ionizing and has the capability to penetrate a wide variety of non-conducting materials such as clothing, paper, cardboard, wood, plastics and ceramics, but is strongly absorbed by polar molecules,

such as water, and reflected by metals (Wallace *et al* 2008). THz imaging and spectroscopy is currently being developed for use in medical imaging (Pickwell-MacPherson and Wallace 2009, Pickwell and Wallace 2006). The first demonstration of biomedical THz imaging (Hu and Nuss 1995) concluded that a distinction could be made between porcine muscle and fat with the hypothesis that the difference in water content of the two materials was responsible for the contrast. Water is one of the main constituents of biological tissue: penetration depths range from typically a few hundred microns in high water content tissues to approximately a centimetre in tissues with a high fat content (Arnone *et al* 1999, Fitzgerald *et al* 2005). Since this first demonstration, the number of reported biomedical studies using THz imaging has increased greatly to include teeth and artificial skin models (Arnone *et al* 1999), healthy skin and basal cell carcinoma both *in vitro* and *in vivo* (Woodward *et al* 2002, Wallace *et al* 2004), excised breast tumours (Fitzgerald *et al* 2006), cortical bone (Stringer *et al* 2005) and burns (Taylor *et al* 2008). Spectroscopic studies in the THz frequency range of excised tissues have shown statistically significant differences between tumours, normal and adipose tissues (Wallace *et al* 2006, Ashworth *et al* 2009).

However, these studies, while demonstrating the capabilities of THz imaging in identifying contrast between biomedical tissues, do not provide conclusive answers with regard to the usefulness of THz radiation as a biomedical imaging tool. There is an apparent need to answer the questions: Does THz spectroscopy allow the measurement of tissue composition and is it sufficiently accurate and sensitive to make a diagnosis of tissue pathologies? Here we investigate the potential of three different analytical methods to detect and specify tissue pathologies at THz wavelengths, as described below.

(i) Linear spectral decomposition method (LSDM). This employs a linear forward model, commonly known as the Beer–Lambert law, based on those used for conventional optical transmittance spectroscopy of media containing different chromophores (Cope 1991), commonly known as the Beer–Lambert law. The concentration of each chromophore present in a material can be estimated from the wavelength-dependent total attenuation of the material using linear matrix inversion. The technique relies on prior knowledge of the absorption coefficient spectra of each of the chromophores in the THz wavelength region.

(ii) Spectrally averaged dielectric coefficient method (SADCM). This is an empirical method based on recent work by Jepsen *et al* (2007) where the real and imaginary dielectric coefficients are averaged over a specific frequency range. By repeating this process for a number of measurements made on samples with different concentrations, empirical equations can be obtained that relate the averaged dielectric coefficients to chromophore concentration.

(iii) Debye relaxation coefficient method (DRCM). Debye theory relates relaxation processes within a material to an alternating external electric field where the material is considered a noninteracting population of dipoles. The method used in this work is empirical and is based on determining the concentration dependence of Debye coefficients from measurements made in phantoms of different compositions. By determining the Debye relaxation coefficients for a number of measurements made on samples with different concentrations, empirical equations can be obtained that relate the Debye relaxation coefficients to the chromophore concentration.

The three analysis approaches are independent of each other: the equations used are given in section 2.3. The same raw reflection data are used in the calculations. Additionally, some prior knowledge is required: for the LSDM, the absorption spectra of pure lipids and water are required, and the SADCM and DRCM both require the spectra of the complex

dielectric coefficient of each emulsion. Transmission spectroscopy measurements were taken to determine these.

In assessing the capability of THz imaging and spectroscopy for medical applications, it is the measurements made in reflection mode which are of greatest relevance. Therefore, the accuracy of the concentration analysis methods was determined on reflection mode emulsion measurements, where the required prior knowledge was determined from transmission mode measurements of pure lipids, water and emulsions.

2. Materials and methods

2.1. Experimental setup and data acquisition

Two stand-alone portable THz systems were used in this study: transmission spectroscopy was performed using the TPI Spectra1000 system and reflection spectroscopy using the TPI Imaga1000 (both produced by TeraView Ltd, Cambridge). All measurements were made at the premises of TeraView Ltd, Cambridge, UK.

2.1.1. Transmission. The spectroscopy unit consists of a sample chamber through which THz pulses are propagated and measured. The chamber is sealed to enable measurements to be made in an atmosphere purged by nitrogen to remove the effects of water vapour from the measured data. Samples for transmission measurements were held in a liquid sample transmission cell which consists of two quartz windows separated by a Teflon spacer. As the cell is designed to be used vertically, rubber rings are placed between the quartz windows and a retaining plate which is tightened to create a water-tight seal. Spacer thicknesses used in this work were 200 μm for water-rich samples and 500 μm for oil-rich samples. Sample measurements were made by measuring the intensity of a THz pulse after it had propagated through the sample chamber and transmission cell containing the sample. A reference measurement was made by measuring the intensity of a THz pulse after it had propagated through the sample chamber and the transmission cell containing no sample or spacer (effectively two pieces of quartz plate held tightly together).

The complex refractive index of the sample is calculated by dividing the measured reference pulse by the measured sample pulse. The amplitude and phase of this calculation are used to calculate the complex dielectric index, from which the absorption coefficient and index of refraction of the sample can be determined. In order to determine the absorption coefficient, it is necessary to approximate the Fresnel coefficients, by assuming that the absorption coefficient is small compared to the refractive index (Jepsen *et al* 2005).

2.1.2. Reflection. Measurements in reflection mode were made in the flatbed geometry set-up of the imaging system, where samples are placed on top of a 2 mm thick z-cut quartz window where the THz beam is focused. The THz pulses are focused onto the samples where the angle of reflection is 30° to the normal. The imaging system was used to collect single-point measurements from the centre of the quartz window, where a THz waveform, approximately 35 ps in length, was acquired for each measured sample. Although an 'imaging' system was used for the reflection measurements, it is the spectroscopic properties of the samples calculated from the reflection pulses which are used, and so subsequently we refer to these reflection data as spectroscopic data. An 'air reference', a measurement of the quartz/air interface at the empty surface of the quartz window, was used in the reflection measurements. All measurements were made at room temperature.

There is also a reflection from the bottom surface of the 2 mm thick quartz window, *R1*, although the main peak of this reflection is not seen in the measured waveform as it does not occur in the measurement time window (0–35 ps). Residual 'ringing' from *R1* does, however, extend into the measurement time window affecting the top sample reflection, *R2*, which should be removed. To eliminate the residual from *R1* a second 2 mm z-cut quartz window is placed on top of the first, creating an 'effective' 4 mm thick quartz window. This has the effect of delaying the arrival of *R2* such that it falls outside the time window (>35 ps) and does not appear in the measured waveform. The measured waveform, therefore, contains only the residual ringing from *R1* and is referred to as 'the background'. The measured data are corrected for by subtracting the background waveform from the reference and sample waveforms. An alternative method for baseline correction has been proposed by Huang *et al* (2009).

The acquired reflection mode waveform represents the impulse function of the sample convolved with the reference waveform, which includes instrumental and environmental aspects such as oscillations in the waveform shape due to atmospheric water. These can be removed by deconvolving the raw waveform with the reference waveform to extract the impulse function. The Fourier transform of this time domain signal gives a frequency domain spectrum (Jepsen *et al* 2007). The refractive index and absorption coefficient were extracted from the raw reflection data following the equations given in reference (Huang *et al* 2009).

2.2. Sample preparation

The water and lipid emulsions used in this study were a commercially available emulsion, Intralipid (Fresenius Kabi, Sweden). Intralipid is a nutritional supplement which is an oil-in-water emulsion of soyabean oil stabilized with 1.2% egg phospholipid with a range of particle sizes between 1 μm and 40 nm (van Staveren *et al* 1991). A series of dilutions of Intralipid with distilled water were produced between 4% and 20% volume lipid concentration. Measurements were also made of pure lipids. The lipids measured were olive oil, safflower oil, beef tallow and lard. The lipids were at least 99% pure with exception of the lard which had been purified manually following a previously described method (vanVeen *et al* 2004). The final purity of the lard could not be determined.

2.3. Theoretical background

2.3.1. LSDM. The LSDM for concentration determination is based on a linear forward model, the Beer–Lambert law, which is typically used for conventional optical transmittance spectroscopy of non-scattering media containing different chromophores (Cope 1991). In this technique, the total absorption of the interrogated medium is considered to be the sum of the absorption contributions of each solute, which are linearly dependent upon their concentration as follows:

$$\mu_{a(\text{total})} = \mu_{a(1)} \cdot c_1 + \mu_{a(2)} \cdot c_2 + \cdots + \mu_{a(n)} \cdot c_n, \tag{1}$$

where $\mu_{a(n)}$ is the absorption coefficient of a compound and c_n is the concentration of a compound. In the THz frequency region, the influence of scattering is assumed to be very small in comparison to that of absorption, meaning that the THz signal amplitude can be taken to be a linear function of μ_a (Smye *et al* 2001, Zhang 2002). The concentration of each solute can be estimated directly using multilinear regression. Calculating concentrations using equation (1) relies on prior information of the frequency dependence of the individual absorption coefficients of the absorbing chromophores, in this case water and lipid. Measurements must

be made at multiple wavelengths using a minimum of n wavelengths for a sample containing n absorbing chromophores.

2.3.2. SADCM. Recent work has suggested that empirical equations can be generated to relate the true concentration of a single solute to the value of the real, ε', and imaginary, ε'', parts of the dielectric coefficients averaged across a frequency range (Jepsen *et al* 2007). These equations are then used as the basis for concentration analysis on future measurements of solutions containing the same solutes. Although this technique does not provide a theoretical model of the physical processes responsible for the concentration dependence of the measured THz signals, it nevertheless incorporates empirical knowledge of the concentration dependence of the absorption coefficient *and* the refractive index through the use of the complex dielectric coefficient. The complex dielectric coefficient of a material, $\widehat{\varepsilon}$, is described by

$$\widehat{\varepsilon}(\omega) = \varepsilon'(\omega) + \mathrm{i}\varepsilon''(\omega), \tag{2}$$

where

$$\varepsilon'(\omega) = n^2(\omega) - \mu_a^2(\omega) \tag{3}$$

and

$$\varepsilon''(\omega) = 2n(\omega)\mu_a(\omega). \tag{4}$$

The implementation of this method for a particular solution relies on establishing the variation with solute concentration of ε' and ε'' over a frequency range. In this work ε' and ε'' are determined from the measured absorption coefficient and refractive index spectra and averaged, at each emulsion concentration, over the frequency range 0.1–1 THz. This frequency range is in keeping with the previously reported study (Jepsen *et al* 2007). The averaged values of ε' and ε'' are then plotted against the lipid concentration from which empirical equations are formed to describe the trend of both ε' and ε'' with lipid concentration. These take the form of linear regressions, such as

$$\begin{aligned} y(\text{real}) &= Ax + B \\ y(\text{imag}) &= Cx + D, \end{aligned} \tag{5}$$

where A, B, C and D are the fitting parameters, y is the averaged value of either ε' or ε'' and x is the concentration of a single solute lipid in the emulsion. Only the concentration of a single solute, in this instance lipid, can be directly determined from the measured data. As the concentrations are calculated as percentages, subtracting the determined concentration of lipid from 100 yields the concentration of water.

The empirical equations form what can be considered to be a 'look-up table' for the determination of emulsion concentrations. ε' and ε'' of an emulsion of unknown concentration are averaged over the frequency range and substituted into the empirical equations, which are then solved for x using the pre-determined fitting parameters. This generates two estimates of the lipid concentration. These are averaged to produce a lipid concentration estimate.

2.3.3. DRCM. Debye theory (Debye 1929) describes the reorientation of molecules which could involve translational and rotational diffusion, hydrogen bond arrangement and structural rearrangement. The Debye relaxation time, the time constant τ, describes the time necessary for $\frac{1}{e}$ of the dipoles to relax to equilibrium after a pulse of EM radiation. For a pure material, multiple Debye-type relaxation processes are possible, and in this work double Debye theory is used:

$$\widehat{\varepsilon}\,(\omega) = \varepsilon_\infty + \frac{\Delta\varepsilon_1}{1 + (\mathrm{j}\omega\tau_1)} + \frac{\Delta\varepsilon_2}{1 + (\mathrm{j}\omega\tau_2)}, \tag{6}$$

which describes two distinct relaxation processes, where $\Delta\varepsilon_1$, $\Delta\varepsilon_2$, τ_1, τ_2 and ε_∞ are the Debye relaxation coefficients. ε_∞ is the real part of the dielectric constant at the high frequency limit. $\Delta\varepsilon = \varepsilon_j - \varepsilon_{j+1}$, where ε_j are intermediate values, occurring at different times, of the dielectric constant, and describe the relative contributions to $\widehat{\varepsilon}$ from the relaxation modes, τ_j. $\Delta\varepsilon_1$ and $\Delta\varepsilon_2$ are equal to $\varepsilon_s - \varepsilon_1$ and $\varepsilon_1 - \varepsilon_2$, respectively, where ε_s, the static dielectric constant, is the real part of the dielectric constant at the low frequency limit. Debye relaxation coefficients were determined from $\widehat{\varepsilon}$ calculated for each emulsion using equation (2). A nonlinear least-squares fitting routine was employed to fit the Debye coefficients to $\widehat{\varepsilon}$ over the frequency range 0.1–1 THz. This frequency range is in keeping with the previously published studies (Pickwell *et al* 2004). ε_s, a frequency invariant material specific parameter, is typically held constant during fitting procedures to improve the stability of the fitting procedure (Pickwell *et al* 2004, Koeberg *et al* 2007). ε_s for the lipid and water emulsions were extrapolated between the ε_s values of pure water and pure lipid, in keeping with other studies, which are typically 78.8 and 2.5, respectively (Pickwell *et al* 2004, Thompson 2005).

As with the SADCM (section 2.3.2) the implementation of this method for concentration analysis relies on establishing the variation of each of the Debye relaxation coefficients as a function of lipid concentration. The Debye relaxation coefficients were determined from the measured absorption coefficient and refractive index spectra of each emulsion and then plotted against the lipid concentration, x. These plots were specific to the trend of the individual Debye relaxation coefficients, and generally took the form of polynomial regressions, such as

$$\Delta\varepsilon_2 = Ax^3 + Bx^2 + Cx + D, \tag{7}$$

where A, B, C and D are fitting parameters, $\Delta\varepsilon_2$ is the emulsion-specific Debye relaxation coefficient and x is the concentration of lipid in the emulsion. Again, only the concentration of a single solute, lipid, can be directly determined from the measured data. As with the SADCM, the determined percentage concentration of lipid was subtracted from 100 to calculate the concentration of water. Again, the empirical equations form a 'look-up-table' for the determination of emulsion concentration. The empirical equations are substituted into equation (6) for each individual Debye relaxation coefficient, which is then solved for the lipid concentration, x, using the $\widehat{\varepsilon}$ of an emulsion of unknown concentration.

2.3.4. Accuracy and resolution. For the concentration analysis methods used here, the accuracy at each discrete concentration was determined by finding the mean difference between the calculated and the true concentration values.

Resolution reflects the smallest detectable change in the determined parameters given the error in the measurement, where a good resolution would correspond to a small value. The resolution in the determined parameters was calculated using

$$\text{resolution} = \sqrt{\text{var}(P)} = \sqrt{(XX')^{-1}\sigma^2}, \tag{8}$$

where P is an array of the variable input parameters, X is the design matrix of the model (X' is its conjugate transpose), which contains the derivative of the models with respect to the input parameters, and σ is the error in the measurement, in this case the standard deviation of the measurement, obtained from repeated measurements.

3. Results and discussions

3.1. Transmission mode 'prior knowledge' of μ_a and n

In this section we present the experimental measurements from samples of pure water and lipids made in transmission mode.

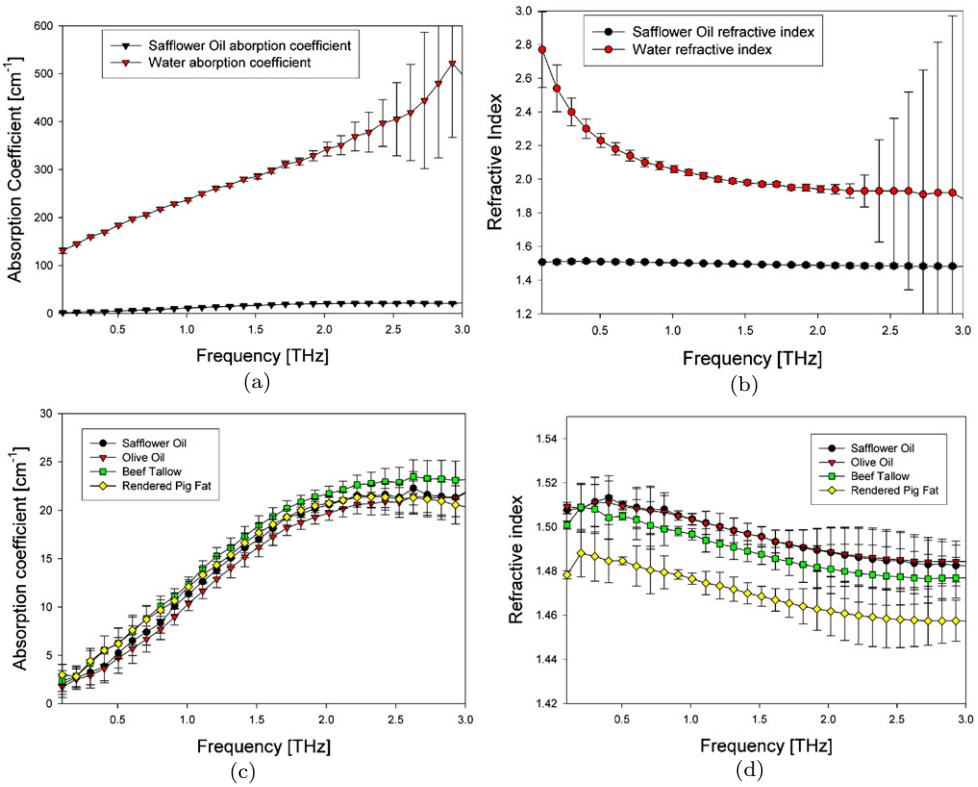


Figure 1. The THz absorption coefficient and index of refraction of pure lipids and water, where (a) and (b) show the absorption coefficients and indices of refraction of pure water and safflower oil to highlight their differing spectral shapes and (c) and (d) show the absorption coefficients and indices of refraction of all the pure lipids. (a) μ_a Safflower oil and pure water, (b) n safflower oil and pure water, (c) μ_a lipids only and (d) n lipids only.

The frequency dependence of the absorption coefficient and refractive index of pure lipids and water are shown in figure 1. The absorption coefficient and refractive index of lipids are smaller than those of water. The lower absorption leads to higher signals and, therefore, greater range of frequencies to be investigated allowing measurements to be made up to a frequency of 3 THz. The absorption coefficient for all lipids increases with a frequency upto around 2 THz. The frequency-dependent absorption coefficient of safflower oil was used as prior knowledge in the LSDM. The absorption coefficients and refractive indices of lipid and water emulsions measured in transmission are illustrated in figure 2 where an increase in both the absorption coefficient and the refractive index with increasing water content of the emulsions is observed.

Values of ε' and ε'' determined from water and lipid emulsions are shown in figure 3 over the frequency range 0.1–1 THz, where there is an increase in both ε' and ε'' with increasing water concentration.

The agreement between the empirical functions and the measured data is illustrated in figure 4(a). The empirical functions describing the variation in ε' and ε'' with lipid concentration are given as

$$\begin{aligned} \varepsilon' &= 5.81 - 0.05x \\ \varepsilon'' &= 3.16 - 0.05x, \end{aligned} \tag{9}$$

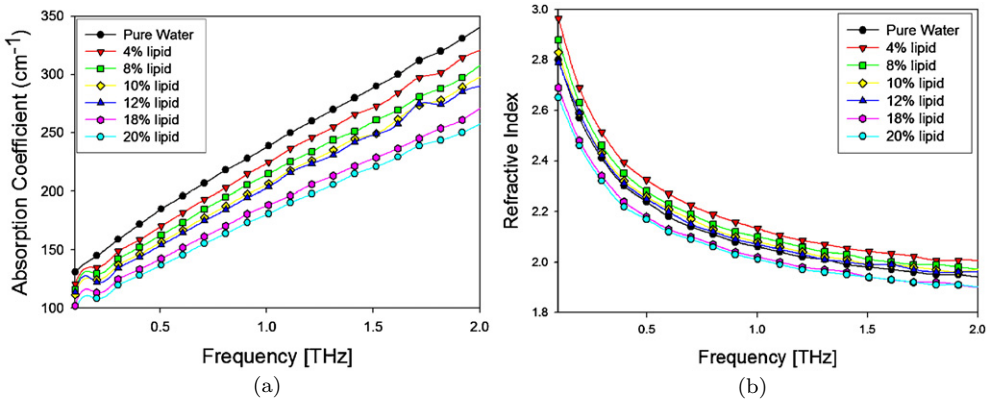


Figure 2. The THz absorption coefficient and index of refraction for different concentrations of lipid in water. The water and lipid emulsions used in this study were a commercially available emulsion, Intralipid (Fresenius Kabi, Sweden), an oil-in-water emulsion of soyabean oil stabilized with 1.2% egg phospholipid: (a) absorption coefficient and (b) refractive index.

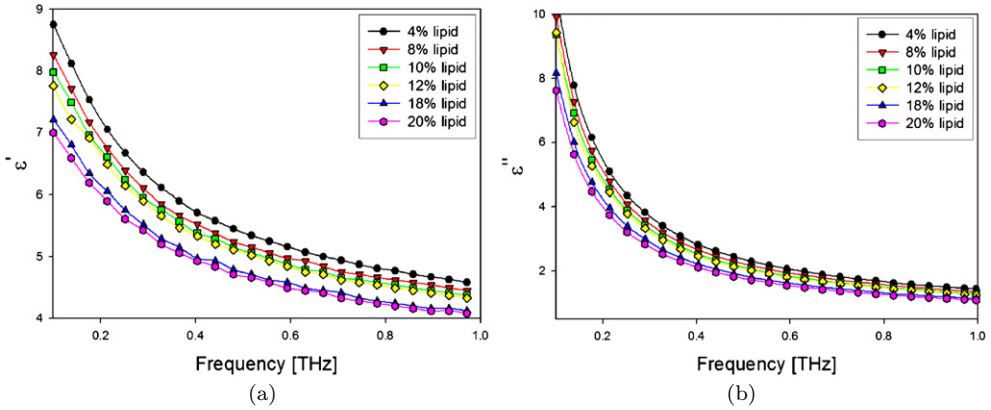
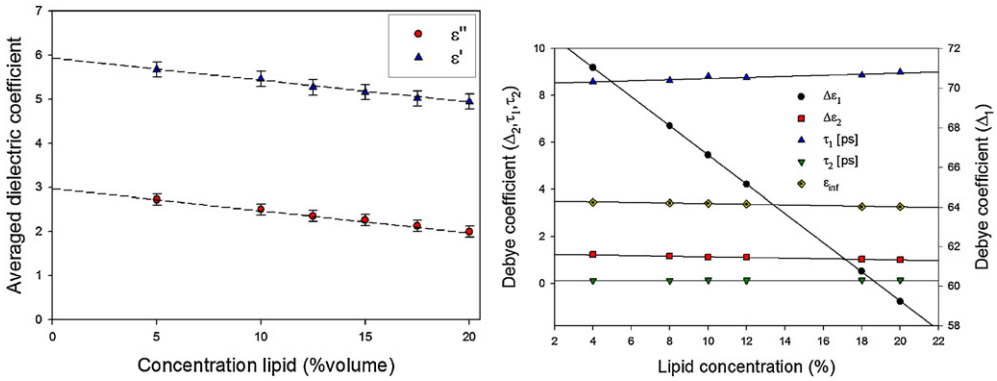


Figure 3. The real and imaginary dielectric coefficients for different concentrations of lipid in water in the THz frequency range. The water and lipid emulsions used in this study were a commercially available emulsion, Intralipid (Fresenius Kabi, Sweden), an oil-in-water emulsion of soyabean oil stabilized with 1.2% egg phospholipid: (a) real and (b) imaginary.

where $x =$ lipid concentration. The variation in the determined Debye relaxation coefficients with respect to lipid concentration are illustrated in figure 4(b). The empirical functions describing the variation in Debye relaxation coefficients with lipid concentration are given as

$$\begin{aligned}
\Delta\varepsilon_1 &= -0.7363x + 73.982 \\
\Delta\varepsilon_2 &= 0.0003x^2 - 0.0213x + 1.3237 \\
\tau_1 &= 0.0234x + 8.5018 \\
\tau_2 &= 0.001x + 0.1243 \\
\varepsilon_{inf} &= -0.013x + 3.5324
\end{aligned} \tag{10}$$

where, again, $x =$ lipid concentration.



(a) (b)

Figure 4. Plot showing (a) the average value of ε' and ε'' in water and lipid emulsions as a function of lipid concentration. The symbols are the averaged dielectric coefficients determined from the data and the dotted lines represent the best fit of the empirical functions. (b) The Debye coefficients measured in transmission in water and lipid emulsions, plotted as a function of lipid concentration. The symbols are the Debye relaxation coefficients determined from the experimental data and the dotted lines represent the best fit of empirical functions: (a) SADCM and (b) DRCM.

Table 1. Standard deviation of 15 measurements of distilled water averaged over the frequency range 0.1–1 THz for reflection measurements.

μ_a (cm^{-1})	n	ε'	ε''
8.6	0.1	0.4	0.2

Table 2. Accuracy and resolution of concentrations determined using LSDM, SADCM and DRCM.

Method	Water Accuracy	Lipid Accuracy	Water Resolution	Lipid Resolution
LSDM	−8.6%	+9.9%	±2.1%	±52.8%
SADCM	−0.4%	+0.3%	±31.5%	±31.5%
DRCM	+3.2%	−3.5%	±10.8%	±10.8%

3.2. Reflection mode concentration analysis

In this section we present the concentrations of water and lipid emulsions as determined using the three concentration analysis methods. The concentrations were determined from the experimental measurements of the emulsions, made in reflection mode. The standard deviations used in the determination of the concentration resolutions are given in table 1 and were derived from 15 measurements of distilled water made over a number of days. The resolution determined in this way gives a more accurate representation of the limitations of spectroscopy and imaging in the THz region. The accuracy and the resolution of these techniques are given in table 2.

3.2.1. LSDM. Figure 5 shows the concentrations determined using the LSDM applied to reflection mode measurements, where a poor correlation between the known concentrations and determined concentrations is observed. The resolution of the water component of the emulsions is good in comparison to the resolution of the lipid component. This is due to the

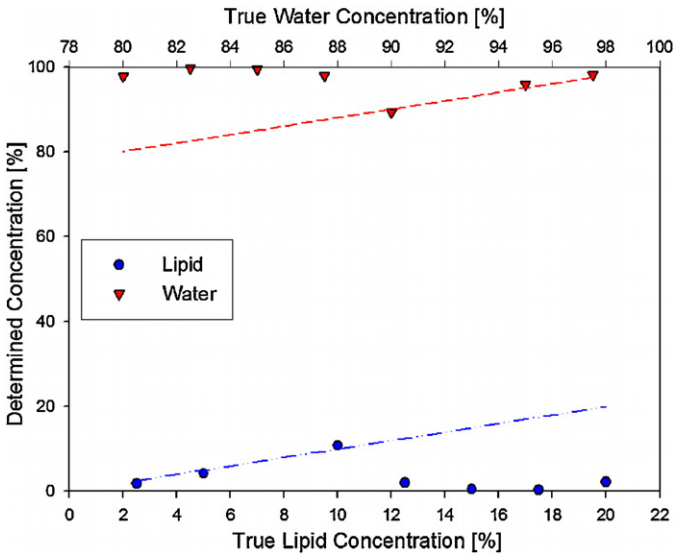


Figure 5. Concentrations of lipid and water determined from measurements of the absorption coefficients of lipid and water emulsions captured in reflection mode using the LSDM. The determined concentrations are plotted against the known concentrations of lipid and water. The line represents the line of unity while the symbols represent the determined concentrations.

large difference in the absorption coefficient which is approximately 20 times greater in water than in lipid at 1 THz. This suggests that it is primarily the water content of the emulsion that THz spectroscopy is sensitive to.

3.2.2. SADCM. The concentrations determined using the SADCM applied to reflection mode measurements are shown in figure 6. From figure 4(a), it can be noted that the gradient of the change of ε' and ε'' with respect to concentration is shallow, resulting in the poor resolution of the method.

3.2.3. DRCM. The concentrations determined using the DRCM applied to reflection mode measurements are shown in figure 7. For the DRCM, the resolutions were determined in a two-step process. First, the resolution in each of the determined Debye coefficients was calculated. These resolutions were then used as the standard deviation in equation (8) in the second step of the process to calculate the overall resolution of the DRCM. The average resolution of this method is still poor in comparison to the concentrations of the emulsions under investigation. For example, the maximum concentration of lipids is 20% which compares to a resolution of 10.8%, 50% of the concentration value.

4. Discussion

4.1. Accuracy of THz reflection spectroscopy

The order of the accuracies provided by the three methods is given below, from best to worst:

(i) SADCM
(ii) DRCM
(iii) LSDM.

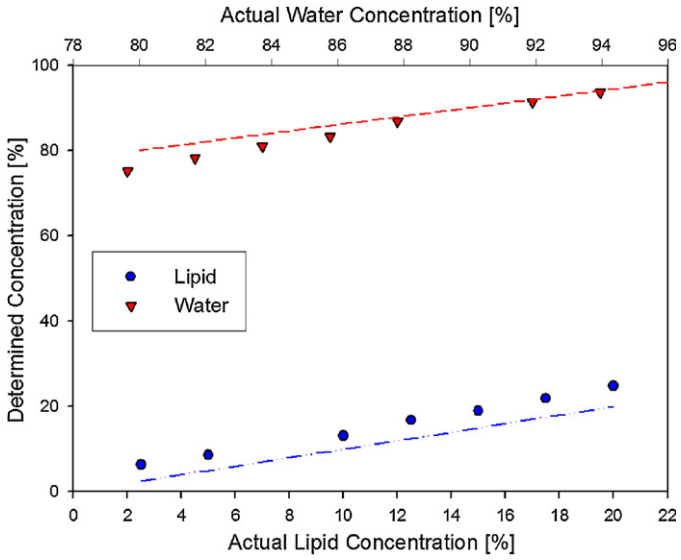


Figure 6. Plot showing the concentrations of water and lipid determined from measurements of ε' and ε'' in water and lipid emulsions using the SADCM. These calculated concentrations are plotted against the known concentrations of water and lipid. The line represents the line of unity while the symbols represent the calculated concentrations.

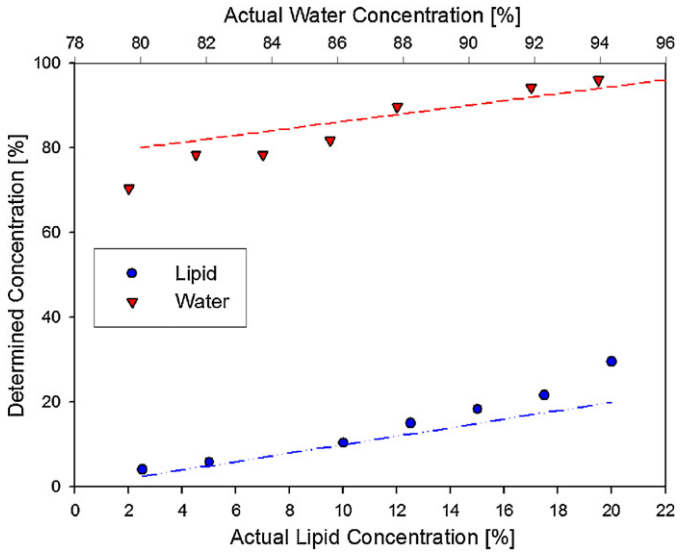


Figure 7. The concentrations of water and lipid calculated using the DRCM analysing the complex dielectric coefficient measured in reflection on water and lipid emulsions. The line represents the line of unity while the symbols represent the calculated concentrations.

LSDM is reliant on only the absorption coefficient in its assessment of concentration and utilizes the absorption coefficients of pure components measured in transmission mode in the analysis of the reflection signals. Small deviations between absorption coefficients measured

in transmission and reflection modes can cause systematic differences. Measurements of the absorption coefficient made in reflection mode have inherently greater systematic error than those made in transmission. In transmission mode, the absorption coefficient is calculated from the amplitude of the waveforms whereas in reflection mode they are derived from the phase components of the measured data and are only accurate if it can be assumed that the incident radiation, at each frequency, can be approximated by a normal incidence plane wave. For the reflection mode imaging system used in this study, beams are incident at an angle of 30°. An EM wave incident at an angle less than 90° suffers diffractive spreading which may result in an error in the absolute value of the measured absorption coefficients and, hence, in the determined concentrations (Bowen *et al* 2004). SADCM and DRCM, incorporate the measured refractive indices of the emulsions as well as the measured absorption coefficients in the concentration calculations, rendering them less sensitive to systematic error than LSDM which uses absorption alone.

4.2. Resolution of THz reflection spectroscopy

The order of the resolutions provided by the three methods is given below, from best to worst:

(i) LSDM
(ii) DRCM
(iii) SADCM.

Resolution is determined by the sensitivity of the method to the measured change in concentration and the standard deviation of the measurement. Since the same standard deviation was used for all calculations of the resolution, the resolution is effectively dependent upon the sensitivity of the method.

The resolutions determined using LSDM were better for the water component which is due to the absorption coefficient of water being much larger than lipid, resulting in the greater sensitivity to water. As the calculation of resolution takes into account the sensitivity for both the water and lipid in the emulsion, a comparatively high sensitivity results in a good resolution. In SADCM, the sensitivity is low, as demonstrated by the shallow gradients of the empirical equations with respect to concentration, which produces a poor resolution. The resolutions produced by SADCM are larger than the actual concentrations. Since any value within the resolution limits has equal statistical significance, it means that the determined concentrations are not uniquely determined. This makes SADCM currently unviable as a THz technique for determining tissue composition. The resolutions for DRCM are better than those of SADCM. This is due to the gradients of the empirical equations of DRCM being less shallow than those of SADCM.

5. Conclusions

LSDM is applicable only to materials in which the frequency dependence of absorption coefficient can be considered to be a linear combination of the component parts. This method requires knowledge of the absorption coefficients of the phantom chromophores as prior knowledge, which are difficult to measure accurately for more complex chromophores such as macromolecules with hydration layers, e.g, macromolecules. The resolutions determined for LSDM are good for large μ_a materials, such as pure water, but resolution is poor for small μ_a materials, such as lipids. SADCM enables accurate determination of concentrations despite the absence of spectral features in the measured dielectric coefficients. If specific signatures could be identified in the dielectric spectrum, then linear decomposition approaches such as

those employed in LSDM may become efficient and enable more complex materials to be analysed. SADCM can only be applied to a maximum of three solutes. In this work only two solutes were analysed, though the method has been demonstrated on three (Jepsen *et al* 2007). In addition it returns poor resolutions, limiting its potential as a tool for tissue spectroscopy. In comparison to LSDM and SADCM, both the accuracy and the resolution for DRCM were reasonable, though the resolution was still poor when compared to the concentration ranges being resolved. Like SADCM, DRCM was applied to the determination of the concentration of only two solutes in this work. This method could be extended to resolve the concentrations of more solutes provided that material-specific variations in the Debye relaxation coefficients can be determined, which are dependent on intermolecular interactions between the solutes.

While the accuracy of the determined concentrations was found to be strongly dependent on the method of analysis used, it was nevertheless shown for all analysis methods that THz spectroscopy is in principle capable of recovering material composition. The greatest accuracy was achieved using the SADCM and DRCM as these are less susceptible to systematic differences between measurements. Nonetheless, the concentrations determined by the LSDM still reflected the general changes in the concentration of the emulsions. This demonstrates significant sensitivity in the measured THz signal to the composition of the emulsions and establishes that, in principle, THz spectroscopy has the potential to analyse tissue composition and specify tissue pathologies.

Although the accuracies for each of the three concentration analysis technique demonstrated the ability of THz spectroscopy to determine composition, the resolutions produced by the three concentration analysis methods were varied. The LSDM has demonstrated a resolution of $\pm 2.1\%$ for the water component of the water and lipid emulsions; however, the resolution of the lipid component was poor. This method is most sensitive to chromophores with large absorption coefficients in the THz wavelength region, such as polar molecules like water and methanol. The sensitivity to concentration of all three concentration analysis methods determined using this method is fixed. In order to improve resolution, therefore, only the standard deviation of the measurement systems may be improved which may be achieved by increasing the signal-to-noise ratio (SNR) of the system.

THz images of tumours show good contrast between diseased and normal tissues (Wallace *et al* 2004, Fitzgerald *et al* 2006) and THz spectroscopy studies have also shown significant differences between tumours and normal tissues (Wallace *et al* 2006, Ashworth *et al* 2009). Here, while we have shown that THz spectroscopy is capable of recovering material composition, our results suggest that it is not yet capable of accurately resolving small enough changes in concentration to enable confident tissue identification. The resolutions are limited by the specifications of the systems used in this study, although continued technological advances in the field which increase the output power and SNR of the systems may enable smaller resolutions to be resolved in the future.

Acknowledgments

CBR acknowledges funding from EPSRC and TeraView.

References

Arnone D D, Ciesla C M, Corchia A, Egusa S, Pepper M, Chamberlain J M, Bezant C, Linfield E H, Clothier R and Khammo N 1999 Applications of terahertz (THz) technology to medical imaging *Proc. SPIE* **3828** 209–19

Ashworth P C, Pickwell-MacPherson Emma, Provenzano E, Pinder S E, Purushotham A D, Pepper M and Wallace V P 2009 Terahertz pulsed spectroscopy of freshly excised human breast cancer *Opt. Express* **17** 12444–54

Bowen J W, Walker G C, Hadjiloucas S and Berry E 2004 The consequences of diffractively spreading beams in ultrafast THz spectroscopy *Joint 29th Int. Conf. on Infrared and Millimeter Waves, 2004 and 12th Int. Conf. on Terahertz Electronics, 2004.* pp 551–2

Cope M 1991 The application of near infrared spectroscopy to non invasive monitoring of cerebral oxygenation in the newborn infant *Phd Thesis* University College London

Debye P 1929 *Polar Molecules* (New York: The Chemical Catalog Company)

Fitzgerald A J, Pickwell E, Wallace V P, Purushotham A, Pinder S, Linan M, Pye R and Ha T 2005 Medical applications of broadband terahertz pulsed radiation *18th Annual Meeting of the IEEE Lasers and Electro-Optics Society, 2005 (LEOS 2005)* pp 120–1

Fitzgerald A J, Wallace V P, Jimenez-Linan M, Bobrow L, Pye R J, Purushotham A D and Arnone D D 2006 Terahertz pulsed imaging of human breast tumors *Radiology* **239** 533–40

Hu B B and Nuss M C 1995 Imaging with terahertz waves *Opt. Lett.* **20** 1716–8

Huang S, Ashworth P C, Kan K W, Chen Y, Wallace V P, Zhang Y and Pickwell-MacPherson E 2009 Improved sample characterization in terahertz reflection imaging and spectroscopy *Opt. Express* **17** 3848–54

Huang S Y, Wang Y X J, Yeung D K W, Ahuja A T, Zhang Y T and Pickwell-MacPherson E 2009 Tissue characterization using terahertz pulsed imaging in reflection geometry *Phys. Med. Biol.* **54** 149–60

Jepsen P U, Møller U and Merbold H 2007 Investigation of aqueous alcohol and sugar solutions with reflection terahertz time-domain spectroscopy *Opt. Express* **15** 14717–37

Jepsen P U and Fischer B M 2005 Dynamic range in terahertz time-domain transmission and reflection spectroscopy *Opt. Lett.* **30** 29–31

Koeberg M, Wu C C, Kim D and Bonn M 2007 Thz dielectric relaxation of ionic liquid:water mixtures *Chem. Phys. Lett.* **439** 60–4

Pickwell-MacPherson E and Wallace V P 2009 Terahertz pulsed imaging—a potential medical imaging modality? *Photodiagnosis Photodynamic Ther.* **6** 128–34

Pickwell E and Wallace V P 2006 Biomedical applications of terahertz technology *Journal Phys. D: Appl. Phys.* **39** R301–10

Pickwell E, Cole B E, Fitzgerald A J, Wallace V P and Pepper M 2004 Simulation of terahertz pulse propagation in biological systems *Appl. Phys. Lett.* **84** 2190–2 **Q3**

Siegel P H 2004 Terahertz technology in biology and medicine *Microw. Theory Tech.* **52** 2438–47

Stringer M R, Lund N D, Foulds A P, Uddin A, Berry E, Miles R E and Davies A G 2005 The analysis of human cortical bone by terahertz time-domain spectroscopy *Phys. Med. Biol.* **50** 3211–9

Smye S W, Chamberlain J M, Fitzgerald A J and Berry E 2001 The interaction between terahertz radiation and biological tissue *Phys. Med. Biol.* **46** R101–12

Taylor Z D, Singh R S, Culjat M O, Suen J Y, Grundfest W S, Lee H and Brown E R 2008 Reflective terahertz imaging of porcine skin burns *Opt. Lett.* **33** 1258–60

Thompson F 2005 Permittivity measurements in solids, powders and liquids *Am. J. Phys.* **73** 787–9

van Staveren H J, Moes C J M, van Marie J, Prahl S A and van Gemert M J C 1991 Light scattering in Intralipid 10% in the wavelength range of 400–1100 nm *Appl. Opt.* **30** 4507–14 **Q4**

vanVeen R L P, Sterenborg H J C M, Pifferi A, Torricelli A and Cubeddu R 2004 Determination of VIS–NIR absorption coefficients of mammalian fat, with time- and spatially resolved diffuse reflectance and transmission spectroscopy *Optical Society of America Biomedical Topical Meetings*

Wallace V P, Fitzgerald A J, Pickwell E, Pye R J, Taday P F, Flanagan N and Ha T 2006 Terahertz pulsed spectroscopy of human basal cell carcinoma *Appl. Spectrosc.* **60** 1127–33

Wallace V P, Fitzgerald A J, Shankar S, Flanagan N, Pye R J, Cluff J and Arnone D D 2004 Terahertz pulsed imaging of basal cell carcinoma *ex vivo* and *in vivo* *British J. Dermatol.* **151** 424–32

Wallace V P, MacPherson E, Zeitler J A and Reid C 2008 Three-dimensional imaging of optically opaque materials using nonionizing terahertz radiation *J. Opt. Soc. Am.* A **25** 3120–33

Woodward R M, Cole B E, Wallace V P, Pye R J, Arnone D D, Linfield E H and Pepper M 2002 Terahertz pulse imaging in reflection geometry of human skin cancer and skin tissue *Phys. Med. Biol.* **47** 3853–63

Zhang X C 2002 Terahertz wave imaging: horizons and hurdles *Phys. Med. Biol.* **47** 3667–77