Chapter from the book *Infrared Spectroscopy - Life and Biomedical Sciences*
Downloaded from: http://www.intechopen.com/books/infrared-spectroscopy-life-and-biomedical-sciences

 Interested in publishing with IntechOpen?
 Contact us at book.department@intechopen.com
Chemometrics of Cells and Tissues Using IR Spectroscopy – Relevance in Biomedical Research

Ranjit Kumar Sahu and Shaul Mordechai*
Department of Physics, Ben-Gurion University of the Negev, Beer-Sheva
Cancer Research Center, Ben-Gurion University of the Negev, Beer-Sheva
Israel

1. Introduction

Biochemical analyses of substances rely upon the ability of techniques to identify qualitatively and quantitatively the components present and are based on physicochemical characteristics as well as chemical nature of substances being detected. While chemical analyses usually depend on reactions of a given substance and can be destructive, spectral studies are usually non-destructive and deal with describing a substance based on properties like absorption or transmission of light (e.g. UV, Visible, Infrared (IR)), light scattering ability, fluorescence/phosphorescence using various optical techniques. Thus, the technique (Fourier transform infrared) spectroscopy has gained prominence in both research and applications in different fields of science. Among the various techniques, IR spectroscopy owing to its lower potency of causing damage compared to X-rays, gamma rays and UV rays (as it is based on weak vibrational energies) has become the technique of choice during chemical analysis of substances. IR spectroscopy can not only provide information about the various components in a complex material but is also unique in its ability to be modified into different kinds of instrumentations based on requirement. The various IR spectroscopy based instruments from a simple IR based spectroscope that helps to obtain the absorbance spectra of a chemical compound to the complex imaging systems that employ computational methods in addition to the technical sophistication are based on a simple principle that every compound or a particular combination of compounds can be described by means of a FTIR (Fourier transform infra red) spectra qualitatively and quantitatively.

The guiding principle of all such analyses lies in the fact that when IR radiation of different wave numbers are simultaneously passed through a sample, specific wave numbers are absorbed based on the vibrations of molecules, creating a unique fingerprint of each sample, from a simple molecule like a protein molecule to a more complex structure like eukaryotic tissues. In spite of the fact that the cells and tissues can be discriminated based on their spectral fingerprints in the mid IR/NIR region, their signatures are the result of contribution from several biological components that at times absorb at similar or overlapping wave
numbers. In order to explain the contribution of different metabolites like carbohydrates, nucleic acids, proteins and lipids in a sample, spectra of pure components are collected and analyzed for specific patterns. Comparison of compound that vary in one or more functional groups aids in the determination and assignment of particular groups and defines the specificity of signature for each along the entire spectral region. Assignment of the exact contribution of each component to the entire spectra therefore makes the method quantitative as well as qualitative.

Figure 1A shows representative spectra of blood fractions obtained by FTIR-MSP (Fourier Transform Infrared Microscopy), mounted on ZnSe slides in transmission mode. It is observed that in spite of the diversity of the source and type of the samples, absorbance of IR occurs at similar wave numbers, implicating that the samples are composed of similar basic substances as mentioned above. Simultaneously it also shows how the variation occurs among these substances and what are the likely principal components of each, contributing to its unique spectra. For example, plasma has less sugar compared to the bacterial cell which has a capsule and thus has more prominent peaks in the region 900-1185. Similarly, the RBCs that lack nucleic acids show diminished peaks while WBCs show clear absorbance peaks between 1185 cm\(^{-1}\) and 1300 cm\(^{-1}\).

![Figure 1A. Amide II normalized average spectra in the region 800-1800 cm\(^{-1}\) of different ZnSe mounted blood components (a-WBC, b-Plasma and c-RBC) analyzed by FTIR-MSP in transmission mode showing few peaks absorbance and the functional groups associated with it.](image)

Thus a qualitative assessment can be made regarding the composition of a sample by looking at its spectra and also defining what type of functional groups could be contributing to absorbance at specific wave numbers.

Several tissues and organisms can show common peaks due to similarity in composition. As shown in Figure 1b, there is absorbance at wave numbers corresponding to components (like carbohydrates, proteins) that are present across a wide range of samples. Owing to large number of data due to the availability of several characteristic wave numbers for individual compounds, mathematical and computational methods are developed that can analyze spectra as per the requirement, providing users the convenience of obtaining the
data in usable and interpretable forms. Advancements in computational techniques have added to the utility of the FTIR based instruments by making spectral calculations rapid and automatic, leading to their application in diverse fields of chemistry and biology for both applied and basic research. The potential of FTIR spectroscopy thus serves not only for routine applications but also as a diagnostic tool where other optical methods become difficult to apply. In view of these developments an ever expanding field of biomedical research based on FTIR based technologies has arisen over the last few decades. With its unique abilities, the technique has been applied mostly in cancer diagnosis and monitoring, microbial identification and drug efficacy evaluations to name a few. The present chapter describes in brief the different aspects of applications of FTIR in biomedicine and their suitability and relevance to biomedical research.

2. Brief historical perspective of FTIR spectroscopy in biomedical research

IR spectroscopy has been in use for chemical analysis for several decades. The recognition of its ability to distinguish between normal and abnormal based on fingerprints of the cells and tissues has been utilized to describe changes at molecular and cellular levels (Yang et al 1995, Malik et al 1996, Schultz et al 1996, Malins et al 1997) opened up a new area in the field of biodiagnostics. Other studies paved way for the examination of the technique as a diagnostic tool for identification of disease especially malignancy (Cohenford et al 1997, Rigas et al 1990, Wang et al 1997). This was followed by expanding the methods to study tissues from several different organs such as skin, cervix, liver among others with an emphasis on cancer identification which would be at par with conventional diagnostic techniques without the negative factors like toxicity to live tissues. Early works focused on defining the tissue variability in terms of IR spectra (Chiriboga et al 1998a,b,c, Wood et al 2004). This led to a large number of studies with different tissues and cell lines with an objective to make the technique relevant to oncology (Fukuyama et al 1999, Diem et al 2000, Gao et al 1999, Yano et al 2000, Malins et al 2003). The potential application in other
biomedical fields was expanded by studying the classification of microorganisms using the technique and using it for identifying pathogens (Kirschner et al 2001, Choosmith et al 2001, Sandt et al 2003, Essendoubi et al 2005). These studies led to the development of the concept of “biomarkers” which are parameters or statistics derived from the spectral data that help to identify or differentiate among samples. The most promising biomarkers have been repeatedly used though developments have led to utilization of more complex spectral information rather than simple ratios. For example the region between 900-1300 cm\(^{-1}\). There were also studies undertaken to identify universally suitable biomarkers which could be used in different cells or tissue types to identify malignancies (Sahu et al 2004b, 2005, Mordechai et al 2004). Thus, there was a parallel development of mathematical and computational methodologies to utilize the spectral data and improve the sensitivity and specificity of the diagnosis and helping clearly defining ambiguous samples or accounting for outliers. These methods include linear discriminant analysis (LDA) (Krafft et al 2006), probabilistic neural networks (PNN) (Mark et al 2004), Principal component analysis (PCA) (Muralikrishna et al 2005) artificial neural networks (ANN) (Zhang et al 2003, Lasch et al 2007) and Discriminant classification functions (DCF) (Sahu et al 2010, Bogomolny et al 2007). The methods like clustering of spectra based on either Wards algorithm using Euclidean distances or Mahalanobis distances have been used to classify tissues and used as a means of developing pseudocolor images during FPA mapping of tissues (Lasch and Naumann, 1998). Similarly while the diagnosis of tissues using FTIR was continuing, the detection of changes in single cells using more sophisticated techniques like synchrotron or SERS were undertaken (Chekun et al 2002). Studies were undertaken to elucidate the contribution of individual components like the nucleic acids which held a great promise as biomarkers (Malins et al 2005, Sahu et al 2008). The works were also supported by inducing transformation in cells using various biological and chemical reagents and studying the spectral manifestations (Ramesh et al 2001, Salman et al 2003, Bogomolny et al 2008). The other significant technical development undertaken was the fiber optic systems that could be used for surface scanning or as probes and these are based on the ATR systems. Parallel research was also carried out to help identify and exclude contaminants in samples (Wood et al 1998, Romeo et al 2003, Sahu et al 2005) or use samples from paraffin embedded systems (Ly et al 2008). The effect of physical phenomenon like Mie scattering and its contribution to the FTIR data was also studied (Kohler et al 2008, Lee et al 2007) with an aim to understand how it can interfere with and influence the basic spectral data. Thus in the last two decades progress has been made in several directions to help realize the full potential of the technique in various biomedical fields making the data both qualitatively and quantitatively relevant. Table I lists a few studies where detection of cancers was carried out using FTIR based instrumentation.

3. Different types of FTIR spectroscopy and their suitability to different fields of biomedical research

Currently the simplest and most familiar IR spectrometers existing consist of a source of IR beam, a sample holder and a detection system to monitor the absorbance. Utilizing several different matrices that are IR transparent (KBr, BaF\(_2\), CaF\(_2\), ZnSe), it is possible to obtain the spectra from samples as diverse as pure compounds such as proteins, lipids, drugs and metabolites to homogeneous preparations such as body fluids, microorganisms, cell lines that can be mounted without destroying their chemical nature. While it is imperative that
<table>
<thead>
<tr>
<th>Authors</th>
<th>Tissue/Organs</th>
<th>Analysis</th>
<th>Region/Wave number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yang et al 1995</td>
<td>Fibrosarcoma</td>
<td>Intensity</td>
<td>1085</td>
</tr>
<tr>
<td>Fujioka et al 2004</td>
<td>Gastric cancer</td>
<td>Discriminant Analysis</td>
<td>925-1660</td>
</tr>
<tr>
<td>Podshyvavlov et al 2005</td>
<td>Cervical cancer</td>
<td>PNN</td>
<td>900-1800</td>
</tr>
<tr>
<td>Krishna et al 2007</td>
<td>Ovarian tissue</td>
<td>Cluster analysis</td>
<td>1540-60 &amp; 1720-1780</td>
</tr>
<tr>
<td>Maziak et al 2007</td>
<td>Oesophagus</td>
<td>Ratios</td>
<td>Several</td>
</tr>
<tr>
<td>Harvey et al 2007</td>
<td>Prostate</td>
<td>FTIR-acoustics</td>
<td></td>
</tr>
<tr>
<td>Argov et al 2002</td>
<td>Colon</td>
<td>ANN</td>
<td>900-1800</td>
</tr>
<tr>
<td>Yano et al 2000</td>
<td>Lung</td>
<td>Ratios</td>
<td>1045,1467</td>
</tr>
<tr>
<td>Wu et al 2001</td>
<td>Oral tissues</td>
<td>Intensity</td>
<td>1745</td>
</tr>
<tr>
<td>Lasch and Naumann 1998</td>
<td>Melanoma</td>
<td>ANN/cluster analysis</td>
<td></td>
</tr>
<tr>
<td>Andrus and Strickland 1998</td>
<td>Lymphoma</td>
<td>Ratio</td>
<td>1020,1121</td>
</tr>
</tbody>
</table>

Table 1. Malignancy in different types of tissue studied by FTIR spectroscopy.

Each sample would require a different approach before being mounted for analysis, the common requirement is removal of water (that interferes with the spectra) through processes like lyophilization or even heating thermo stable samples in an oven. The samples are at times dissolved in D$_2$O to overcome the effect of water especially where they are aqueous soluble. Such systems are used to study interactions between biological molecules such as proteins and nucleic acids with ligands like toxins and antibiotics (Pratibha and Malathi, 2008, Bourassa et al 2011).

A slightly different approach is required for cells and microbes where a homogeneous material is available (e.g., cell lines, bacteria). Such samples require removal of extraneous materials that can confound the spectra and are usually subjected to washes by normal saline before being dried for spectral acquisition. For example, bacterial cells from cultures in log phase are harvested after treatment with chloroform and the media removed by centrifugation. Subsequently they are washed several times in normal saline (without phosphates) and mounted on discs or any IR transparent support for spectra collection (Sahu et al 2006b).

When dealing with eukaryotic cells like cell lines, the cells are harvested and washed in a buffered solution several times to get rid of the media components and finally washed in normal saline before removing the excess water by a process of air drying. Similarly, different components of blood can be separated and fractions of plasma directly dried on the support. For the cellular components of blood or fluid biopsies like fine needle extracts by different biophysical techniques can be used for obtaining a given population of cells such as RBC, WBCs, monocytes or T cells and B cells which after purification are treated similar to cell lines for spectroscopic measurements. Most of these samples can be measured with routine spectrometers.

The approach becomes more complex when dealing with spectroscopy of tissues. In case of a relatively homogeneous tissue like liver, spectra can more or less be obtained from any region of the biopsy while tissues like cervical tissues that have clearly differentiated zones would need measurements at precise locations to alleviate spectral variation due to location
effects (Chiriboga et al 1998a). However these samples necessitate a microscopic evaluation before measurement. Colonic biopsies that display a more or less uniform pattern in the cross section are measured on the circles of crypts for diagnosis of malignancies. However the longitudinal sections of the crypt require a more defined location for measurement and invariably depend on the utilization of microscope. More complicated measurements of tissues where several different kind of cells are required to be measured depend on focal plane array detectors where each pixel of the measured area can be represented by a spectra. The requirement of methodologies for reconstruction of a pseudo image based on spectral characteristics become the norm and essential feature of such measurements.

4. FTIR instrumentation and their applications in biomedicine

The simplest FTIR spectrometers have been most widely used to study cells and plasma samples. When a sample’s spectra are obtained, they are normalized, averaged and baseline corrected. The usual practice is to undertake a baseline correction in the spectral region of interest using the rubber band form and locating the two extreme points of the region of interest. The cut spectra are then normalized again to the highest peak or the area. The different methods of baseline correction in specified regions would greatly alter the results. Hence application of similar data processing on all spectra being considered in a study becomes essential. The intensities at selected wave numbers are then used to define ratios or biomarkers that can define the criterion being sought for the diagnosis. Extracting intensity ratios is preferred than band intensity, since it yields a dimensionless quantity which is mostly independent on the exact normalization procedure used (such as Min-Max or vector normalization). Other than the intensities, many studies report shifts in peak wave numbers which can provide additional valuable information. However instruments that can measure only a few wave numbers can also be used when more sophisticated instrumentations are not available. For example measuring the band intensities of the CH$_2$ and CH$_3$ vibrations between 2800-3200 cm$^{-1}$ can provide information about the status of the tissue (Sahu et al 2006a,2005). Similarly various other parameters have been routinely used in diagnosis of diseases and cancer. The ratio of Amide I /Amide II quantified using the integrated absorption of Amide I and Amide II (1750-1590 cm$^{-1}$, 1590-1480 cm$^{-1}$) was one of the first ratios found suitable for diagnosis and is an indicator for the DNA absorbance variation (Liquier and Talliander,1996, Benedetti et al., 1997; Gasparri & Muzio, 2003). RBCs (red blood cells) from humans that lack nuclei posses a ratio of unity for this parameter and maybe used as a reference when understanding nuclear changes using these wave number regions.

Similarly quantification of phosphate metabolites calculated by measuring the integrated area of phosphate symmetric (990-1145 cm$^{-1}$) and antisymmetric (1190-1275 cm$^{-1}$) bands have been used to understand the stages of the cell (Yang at al. 1995). The most promising phosphate absorbance that has been used in many studies is the RNA/DNA ratio dependent on the phosphate absorbance arising from the symmetric vibrations of DNA and RNA namely at 1020 and 1121 cm$^{-1}$ respectively. Several other band intensities have been used individually or in combination as listed in table 1 for evaluating cells, tissues and biopsies. The band at 1045 cm$^{-1}$ is attributed to the vibrational frequency of -CH$_2$OH groups and the C-O stretching frequencies coupled with C-O bending frequencies of the C-OH groups of carbohydrates (including glucose, glycogen, etc.) as well as the capsular
carbohydrates of bacteria. The ratio of the integrated absorbance at 1045 cm\(^{-1}\) to that at 1545 cm\(^{-1}\) provides an estimate of the carbohydrate absorption (Parker, 1971). When a normalization is made to the amide II band the intensities at this wavenumber can be assumed to directly correlate with the amount of carbohydrates and is similar to normalizing the amount to protein in cells and tissues using chemical analyses.

5. Protocols for sample preparation

As the sources of samples for FTIR spectroscopy can be different, they need special preparations adopting different kinds of procedures depending on the samples.

5.1 FTIR spectroscopy of homogeneous materials

5.1.1 Drug macromolecules interactions

Studies of drug interactions by FTIR spectroscopy are easily carried out in systems using D\(_2\)O. Various approaches are possible. The drug and protein/nucleic acid interaction is carried out under solution conditions at predetermined stoichiometry and the mixture is then added to a film from which spectral data are obtained. The individual components are also used as references or controls. The difference spectra between the control and with the ligands is sued to quantify the amount of bound ligands or study the changes in the protein secondary structure by monitoring the shifts in peak intensities or variation in peak intensities of different functional groups (Bourasssa et al 2011). FTIR-ATR system can also be used for studying the interactions using a dialysis system. (Krasteva et al 2006, Kumar & Barth (2011).

5.1.2 Microbial cultures

Bacterial cultures growing in the log phase are fixed with formaldehyde (final concentration 0.25%) and washed three times with saline containing 0.25% formaldehyde and once with saline only. The pellet is resuspended with saline to an OD\(_{450}\) = 0.3 (about 20 ml). One microliter of the suspension is spotted on Zinc-Selenium slides and air dried for 6 hours to remove any water in the sample under a laminar air flow chamber. Microscopic FTIR measurements are made in transmission mode using the FTIR microscope IReScope II with a mercury-cadmium-telluride (MCT) detector, coupled to the FTIR spectrometer (Bruker Equinox model 55/S, OPUS software). Absorbance is measured from an area of 100 nm diameter by setting the slit to 100 microns. Regions of thickness of about 10 microns (as seen from the ADC values) are selected and 128 scans co added for each spectrum. For each sample at least five spots are measured and the average spectra calculated. Similar procedure can be adopted for yeast cells. In case of fungi which have a tendency to grow as filaments (hyphae), the filaments are harvested and briefly tweezed or torn to get uniform untangled mass as observed under the microscope. These are then washed in water or saline before mounting on the slides. These can also be measured using ATR systems as they are highly adherent and a spread sample is likely to have non uniformity in thickness owing to the clumping tendency of the filaments (Salman et al 2011).

5.1.3 Cell lines and transformed cultures

Cell lines are cultured under suitable conditions (e.g. 37°C in RPMI medium supplemented with 10% of newborn calf serum (NBCS) and the antibiotics penicillin, streptomycin) and
passed to obtain confluent cells. Adherent cells are harvested using a cell lifter while cells in suspension are taken directly. The cell suspensions is centrifuged for 5 minutes at 3000 rpm and the pellet washed in phosphate buffered saline several times (3-5 times). After the washing, the cells are washed at least two times in normal saline and resuspended in saline such that the cells density is about 1x10^6/ml. One microlitre of this suspension is spotted on a ZnSe window and allowed to air dry in a laminar air flow chamber for several hours. A microlitre of the last supernatant before suspension may also be spotted to be used as a reference for any unwanted materials. Spectra from this sample should ideally be similar to the background spectra. Cells and microbes like fungus may also be grown on IR transparent matrices when they are analyzed by ATR or in reflectance mode. Precautions are however required to avoid artifacts due to very thin layers of drying of cells (Mourant et al 2003a).

5.1.4 Samples from body fluids such as blood or urine

Body fluids usually contain cells and non cellular components, tissue debris and a fluid or matrix. When the cells are the material of interest, they are isolated into pure or homogeneous forms using various separation techniques that utilize the different sedimentation coefficient of the cells and including Fluorescence activated cell sorter (FACS) when required to isolate a pure population based on their cell surface markers. In case of blood samples, the collection is made in heparin or EDTA to prevent the clotting. Then they are separated by a histopaque or Ficoll gradient. The cells are then washed in PBS followed by normal saline as for cultured cells. For spectra of other components such as plasma from blood, the plasma or sera obtained from blood is directly spotted on the ZnSe windows without further processing. However, if the liquid is very viscous, it can lead to supersaturated spectra. To avoid this, the fluid is spread by dragging the drop in several directions with the tip and spreading it. Study of other body fluids like urine may require filtration of the liquid to remove debris and its concentration to reduce the amount of water before it is loaded on matrices. In case of study of the cellular components, a prior separation of different cell components maybe required.

5.2 FTIR spectroscopy of tissue samples

Most of the diagnostic potential of FTIR spectroscopy has been evaluated in tissues that relate to carcinogenesis with an objective to diagnose malignancy or premalignancy. Thus, conventionally the material available was formalin fixed, paraffin embedded tissues. In most of these cases, the tissue is sliced to a suitable thickness (usually 10 microns), processed to remove the embedding material and mounted on a slide (usually a ZnSe slide) before it is used for measurements. Such slides are stored in histology frames dipped in 70% ethanol to prevent contamination. A consecutive section that is congruent to the section being measured is usually used for H&E staining. Thus, both the optical as well as the FTIR measuring capabilities of the microscope are used to pin point the exact location of measurement. For example, in tissues like cervical epithelium, where there is a decrease in glycogen level from the superficial layer to the basal layer (Chiriboga et al 1998a), and also the reduction of glycogen is an important marker for CIN or cervical cancer, it becomes important to measure at defined regions, namely the intermediate layers of the tissue to accurately be able to classify the stage of CIN (Figure 3a) (Mark et al 2004). In case of colonic epithelium, the tissue can
contain regions where the crypts are seen in a transverse or longitudinal section. The measurements can be made on either type of section. However, the transverse sections are usually measured for conventional diagnosis and grading for cancer or premalignancy (Argov et al 2002, Argov et al 2004). Measurements along the crypt height have been employed to understand metabolic activity in terms of spectral changes (Salman et al 2004) or to define abnormal proliferation in the epithelium (Sahu et al 2004a, Sahu et al 2010). In these cases, a region of the slide free from contaminants is used for the background. The histology of the tissue is evaluated on the complementary slide which is stained with H&E. Regions are selected that indicate the required stage of disease and then measured by locating the identical spots on the ZnSe slide. At least five measurements on each sample are carried out. The morphology of the colonic epithelium is well defined and it is easy to locate the crypts even while viewing under the optical microscope. Thus, it is easier to locate the crypts and measure them. Usually in these tissues, measurements are made by adjusting the slit to 120 microns that results in an area of 100 micron diameter. As most formalin fixed tissues are paraffin embedded and processed, the spectroscopy in the higher region can also be used to monitor contaminants like blood and paraffin or any residual cleansing reagent like xylol (Sahu et al 2005). Both the transverse sections of the crypt, where the entire crypt would fit into the area of measurement or several locations along the crypt in the longitudinal section can be measured based on the requirement of the studies (Argov et al 2002, Argov et al 2004, Sahu et al 2004a, Salman et al 2004). In case of measurement along the length of the crypt the replication is carried out by measuring several crypts at similar distance form the base or apex unlike the cross section where several crypt circles are measured. For more complex tissues such as skin, where different kinds of abnormalities occur (both benign and malignant) the exact location of the sample measurement area requires the involvement of an expert dermatologist/pathologist to pin point and mark the areas on the complementary slides (Hammody et al 2007). Advancement in technologies have made it possible to automatically measure entire regions of the biopsy using the FPA detector systems and the spectra corresponding to every pixel of the measured area can be obtained. This type of measurement is slow though automated and helps to map entire regions. However the study on a complementary slide to demarcate the exact location of histological entities is still essential to precisely correlate the spectra. The processing is similar to microscopy but owing to a large amount of acquired data, automated computational methods become necessary.

5.3 Cellular transformations

Cellular transformations are used to understand spectral changes happening due to biological changes in cells and tissues, often inducing controlled changes in cell lines using genes or viruses. Cells grown in tissue culture plates are treated with polybrene (a cationic polymer required for neutralizing the negative charge of the cell membrane) for 24 h before infection with the virus. Free polybrene is removed and incubated with the high titer infecting virus stock at 37°C for 2 h. The unabsorbed virus particles are removed by washing the cells in fresh warm medium and fresh medium is added. Several wells are used in a study including a control group passed through similar conditions without the infecting virus. After various stages of progression the cells are examined for the appearance of malignant transformation and cells from sets of wells harvested at different time intervals used for spectra acquisition. Cells can also be grown on soft agar or stained to confirm their malignant transformation.
6. Spectra acquisition and data processing

Spectra can be obtained either in transmission or reflective mode or both from a FTIR microscope (Argov et al 2002, Chang et al 2003). Samples loaded on a mounting material such as a ZnSe slide are observed under the microscope for their uniformity and thickness and representative regions are selected. Prior to measuring the samples a background measurement is made in a region free of any samples or reference material. This spectrum is saved as a background (Figure 2a). The data processor automatically subtracts these

![Graph](image)

A. Initial absorbance spectrum collected in transmission mode (a), background (b) and background subtracted and rubber band baseline corrected spectrum (c). B. Spectra in the region 800-1800 cm\(^{-1}\) after cutting from the entire spectrum, baseline correction and normalization to amide I peak (a) or amide II peak (b). C. Area normalization of the spectrum after cutting the region between 800-1800 cm\(^{-1}\) preceded by baseline correction and followed by offset correction. The vertical lines denote intensities at 1545 and 1045 cm\(^{-1}\) which can be used for quantification of carbohydrates. D. Deconvolution of the spectra to obtain intensity at various peaks using Gaussian fit of the spectra. (E.) Second derivative of the spectra. (F) Selected region in the second derivative spectra where the intensity at a selected wavenumber is calculated by taking the intensities at the minimum (a) and the adjacent maximum (b).

Fig. 2. Spectral acquisition and processing.
background spectra when further measurements are made. The reference samples if any are then measured. The thickness of the sample is usually reflected in the ADC rates. A ADC value of less than 3000 usually denotes a thickness of less than 10 microns and often leads to noisy spectra. On the other hand regions with a very high ADC value of (>6000) can give rise to supersaturated spectra. Once a set of reference spectra are obtained and they display required characteristics, the samples are measured. Ideally 3-5 spots are measured on each spotted sample.

The background corrected spectra is then baseline corrected for the entire region using a rubber band baseline function (Figure 2A). Selected spectral regions of interest are then separately cut from the entire spectrum and a second baseline correction is made (e.g. region between 800-1800 cm\(^{-1}\)). This is followed by a normalization of the spectra (Fig.2B,C). The different spectra are then averaged and the average spectra used to represent a sample. Intensities at different wavenumbers obtained from this processed spectra are the inputs for various mathematical analyses. For example, the ratio of intensities at 1545 and 1045 cm\(^{-1}\) (Figure 2C) or peak areas after deconvolution (Figure 2D).

Usually a region with normal morphology is taken as a reference or control while identifying the malignancy in a tissue to avoid any heterogeneity due to individual variations while comparing normal and abnormal tissues (Argov et al 2002). With recognition of the fact that normal tissues may still display abnormal FTIR spectra it becomes important to precisely define a normal tissue or biological entity which in itself could lead to better understand dynamics of spectral variation during disease progression (Sahu et al 2010).

Second derivatives and higher order derivatives of spectra are used to avoid the errors creeping in due to bias originating from baseline selection methods (Figure 2E). In this case when the intensities are measured, the minima correspond to the maxima (peak intensity) in the original spectra. Sum of the value at the minima with the value at the nearest maxima is taken to indicate the intensity at the particular wavenumber (Fig 2F).

### 6.1 Spectral acquisition using ATR/FEWS systems

ATR measurements can be done using Bruker FT-IR Tensor 27 Spectrometer equipped with a liquid nitrogen cooled mercury-cadmium-telluride (MCT) detector and coupled with Horizontal Attenuated Total Reflectance Accessory (HATR (Horizontal Attenuated Total Reflectance), PIKE technologies Inc.) systems (Bogomonly et al 2009). The accessory is connected with a nitrogen reservoir, which enables to preserve dryness of the sample and maintaining an inert atmosphere. The design employs a pair of transfer optics to direct the infrared beam to one end of IR transmitting crystal. A similar pair of flat mirrors directs the beam from the other end of the ATR crystal to the spectrometer detector. The ATR crystal made up of ZnSe is of a trapezoid shape with its thickness suitable to produce optimum performance. To reduce variance each sample is measured several times (at least five) and the ADC rates are empirically chosen between 4000-5000 counts/sec.

### 7. Advances in data analysis

Unique spectral fingerprints of biological entities in the mid IR region are the manifestation of several components absorbing at different wavenumbers, with overlap. Thus, both the
quantity and type of individual components can alter the fingerprint. As shown in Figure 3, during carcinogenesis, in both colon and cervical tissues, there is a depletion of carbohydrates likely due to increased metabolism. However, the disappearance of carbohydrates evident from vanishing of the triads between 900-1185 cm\(^{-1}\) in cervical tissues is slightly different than in colonic tissues. This is likely because glycogen is known to accumulate in cervical tissues, increasing in concentration from the basal to superficial layer, the absorbance associated with colonic tissues is more likely to be from glycoproteins than pure glycogen. Thus, though similar functional groups may contribute to the absorbance, they can manifest as spectral variations. This also necessitates that each tissue or cell type is investigated independently though common biomarkers are used. Such type of differences and the gradual variation at specific wave numbers due to transition of tissues or cells from one type to another requires that contribution of individual metabolites like carbohydrates, nucleic acids, proteins is clearly evaluated.

Fig. 3. Baseline corrected spectra from (top) normal and cancerous colonic tissues and (bottom) normal and cancerous cervical tissues, indicating changes in the region 900-1200 cm\(^{-1}\). The circles in the histological sections depict the measurement sites. Note that spectra in the upper panel are further baseline corrected by using a rubber band baseline in the region 900-1200 cm\(^{-1}\) while in the lower panel the data presented are after normalization to the amide II peak (not shown). Note also the similarity in the trends in peak intensity on carcinogenesis irrespective of tissue origin.

Figure 4a shows the variation in the integrated intensity in colonic biopsies with different diagnostic features. It is noted that while both normal and hyperplastic samples show
similar quantitative traits, the other biopsies with varying degrees of malignancy or premalignancy display a decrease in the levels. It is also noted that the samples with worst prognosis show similar values (For example C2 grade cancer and Severe level of hyperplasia have similar value). This indicates that systems with similar spectral features pertain to a class of conditions with similar outcomes. The feature can also be used to monitor the progression of the disease over time, indicating its possible potential as a biomarker (Figure 4b).

However the two groups of tissues need separate independent analysis, i.e. the samples of cancerous biopsies and hyperplastic biopsies when dealing from a clinical perspective. To further be able to differentiate between these two subgroups additional biomarkers become necessary. The availability of several biomarkers, suitable for disease diagnosis can complicate the selection though they can increase the sensitivity of the technique. Therefore other than using simple ratios, cluster analysis using several biomarkers or spectral data in entire regions can be undertaken. These help to minimize the false negatives or positives. Inclusion of mathematical and computational methods to analyze suitable wavenumbers for diagnosis focused on the differences between normal and abnormal tissues can utilize artificial neural networks and set up data bases that are used as a reference. Usually a part of the study sample is used to train the system before the blind samples are analyzed. Setting up of reference data base with spectra from clearly identified histopathological systems is a primary step in setting up of a good diagnostic software.

The potential of FTIR increased several fold by the combination of computational methods can thus also be used to overcome spatial and temporal variation in samples. The setting up of such database becomes more crucial when microorganisms (that have a tendency to mutate and change rapidly) are being studied as the older database can be used to monitor such evolution by studying spectral variations. Cluster analysis where different groups are separated by the distance proportional to their heterogeneity is another way to study the evolutionary relation between species and subspecies of microorganism. However clustering may not be sensitive enough to discriminate among closely related samples. Thus,
more methods of analyses are resorted to like ANNs (Goodacre et al 1996). Artificial neural networks make it possible to examine samples over time by setting up reference data bases. These systems often work on the principal of classifying the tissue in a binary progression mode as depicted in Figure 5a. The final results are displayed as a confusion matrix where the probability of classifying a particular biopsy into one of the diagnostic group is expressed as a percentage.

![Diagram of FTIR Sample Classifier](image)

Fig. 5. A model displaying a possible methodology to classify cervical biopsies using an ANN.

Data analysis involving FTIR spectra focuses on utilization of intensities or integrated absorbance from wave number/ wave number regions or their various combinations that result in separation of the different classes of samples under study. Often diagnosis between normal and abnormal tissues is carried out by monitoring absorbance at selected wave numbers after routine mathematical manipulation of the spectra. (Sahu et al 2004, Mordechai et al 2004). Later entire regions of spectra or their derivatives were used to classify tissues or samples (supervised or unsupervised) into clusters to determine their hierarchy or their relation with one another. These types of data conventionally presented as clusters were later used in advanced methods like FPA, to organize areas with similar into pseudocolor maps to establish patterns of tissues from FTIR spectra.

Cluster analysis based on the Ward’s algorithm separates samples by the distance proportional to their heterogeneity and difference spectra of colonic crypts has successfully been used their classification (Sahu et al 2004). Similarly the cluster analysis has been used for classification of microorganisms (Sahu et al 2006). Figure 5b displays a schematic diagram of how different biopsies are classified and the closely related conditions tend to group together. Cluster analysis of FEWS spectra of human skin samples using a chemical factor analysis has been shown to differentiate melanoma from basaloma (Sukuta and Bruch 1999).

The data analysis methods like DCF are more useful when classifying systems using morphological features like crypts compared to the others (Sahu et al 2010) as each of the potential biomarkers is used with its weighed contribution which helps to discriminate between normal and abnormal biopsies by representing an adequate quantitative follow up of transformations versus time. DCF can therefore used to classify biopsies using a
classification score that is a linear combination of several potential biomarkers taking advantage of the variation at several wave numbers. Usually the normal or abnormal class is given a final value of 100 and its reverse condition the value of zero by using a combination of weighted coefficients. The different intermediary stages lie within these two extreme values (Hammody et al 2007, Bogomonly et al 2009). One or more of the statistical and computational methods can be resorted to classify samples when an ambiguity is encountered.

Classification of not only the normalized spectral intensities or spectral regions but also the spectral derivatives or difference spectra may be undertaken while performing cluster analysis. Confusion matrices are obtained as a result of ANN and PNN analysis which help to classify tissues or biopsies based on their probability of falling into a particular class. The LDA and PCA techniques classify biopsies into several groups presenting them in normal graphical formats and assigning the grade. The DCF analysis improves discrimination between different stages by using a weighted value for each biomarker used and enables proper assignment during studies involving progression of changes in cells and tissues (Hammody et al 2007, Bogomonly et al 2009).

8. Recent trends and future perspectives

Though basic research has been extensively carried out including pilot scale experimentation of clinical trials in using FTIR measurements for diseases diagnosis, the potential has not been practically exploited due to inability of FTIR based diagnosis to be an independent technique for classification of diseased tissues. The role of a pathologist has been indispensable and pivotal in the preliminary process of sample selection. Advances made through development of FPA based techniques partially overcome this requirement where automated identification of diseased regions in tissues using programs like cluster analysis or ANN/PNN provide pseudocolored images depicting tissue morphology. The rapidity of the technique is however compromised in these cases. Moreover, while a pathologist could quickly look at areas of interest, the automation mandatorily examines the entire section, making it a time consuming affair. In case of complex material like melanoma
and different grades of nevi, the process still needs supporting techniques like immuno
histochemistry/histology for verifications before an FTIR measurement is made. The
complex dynamics of epithelial tissues that are most prone to carcinogenesis, vary from one
organ to another. Thus, clear knowledge of the type of metabolic pattern variation in these
tissues is a prerequisite for both acquisition and interpretation of the spectral data. This
makes the technique interdisciplinary and involves people with extra specializations
requiring special training. However owing to the simplicity in data acquisition and
automation with most post spectral processing, these limitations may be overcome easily.

An objective and quantitative method like FTIR spectroscopy becomes crucial when
metabolomics of cells and tissues are desired without requirement of different chemical
analyses, as it simultaneously can monitor different metabolites. It has the ability to
distinguish between microorganisms depending on their biochemical variation and this has
been demonstrated in both bacteria and fungi (Sahu et al 2006b, Salman et al 2011).
Similarly, minute changes can occur between the different stages of cancer or malignancy
which are difficult to detect even with histochemical analyses. Under these circumstances,
FTIR becomes handy as it distinguishes among different grades of premalignant and
malignant tissues based on markers like carbohydrates which may not be easily stained in
normal laboratory practices (Mark et al 2005). Observation of metabolite variation along
tissues reflecting the dynamic nature through FTIR spectroscopy has been one of the most
important findings (Chiriboga et al 1998a, Salman et al 2004, Sahu et al 2010) that has
implications for cancer diagnosis. For example abnormal crypt proliferation can detect in
colic tissues to re-evaluate resection margins (Sahu et al 2010) consequently decreasing
the number of surgical interventions. The technique becomes an inexpensive alternative
method to follow disease progression or regression quantitatively over a period of time
(Zelig et al 2011). Similarly, the growth patterns of microbes are being studied using FTIR
spectroscopy. The technique has also become a method of choice to monitor drug effects on
cells and tissues as it can easily monitor changes occurring due to onset of processes like
apoptosis and necrosis (Zelig et al 2009).

In addition to these abilities, the availability of a qualitative evaluation along with
quantization adds in determining the status of the tissues. For example, as shown in figure
6, the values of the ratio of integrated absorbance of several crypts in colonic biopsies shows
two different patterns. The biopsies of patients diagnosed with polyposis show a low ratio
of 1 or less for many crypts while biopsies without any polyposis have most crypts
displaying a the ratio greater than 1. While conventional statistical analysis would look at
values like the average for determining the probability of the biopsy being normal or
abnormal, the present system allows the pathologist to examine each biopsy with the
privilege of being able to designate the biopsy as abnormal based on the values of more than
one crypt going with the medical practice of assigning the worst prognosis. This is an easier
approach where spectral data from the biopsy cannot be analyzed using more complicated
methods like DCF analysis. Thus, the spectral analysis can be tailor made to suit the
availability of qualified personnel.

The fiber optic based IR probes are being increasingly examined for utilization as in vivo
detection systems as these are non toxic. Several reports on fiber-optic based IR sensors for
diagnosis of malignancies reflect the trend of incorporation of this technique in future
clinical practices (Mackanos and Contag 2010). This in combination with light based
colonoscopes, colposcopes or other endoscopic tools maybe used to detect malignant or abnormal tissues in body cavities. While the optical systems could help to pin point locations, the FTIR system could provide a quantitative evaluation of the abnormal tissue in vivo, without the requirement of the pathologist and help the surgeon make real time decisions. This is especially important as histologically normal regions can be diseased when evaluated by FTIR spectroscopy (Argov et al 2004, Sahu et al 2004, 2010). For example, IR probes mounted on colposcopes may not only determine the areas of malignancy in the cervix but also possibly be able to determine the stage of the CIN. Similarly, examination of body fluids or other material for microbial infections without requiring long culture time can be rapidly done if detection systems are developed where the type of changes associated with different kinds of microbes are clearly defined and established (Bombalska et al 2011, Maquelin et al 2003). The ability of FTIR to detect biochemical changes in organs other than the one affected to possibly use these as indicator of the health status of patients has been another growing area of interest, mainly because these use materials like hairs and nails that are dispensable or blood samples that are easily obtained compared to biopsies (Lyman et al 2005, Khanmohammadi et al 2007). This type of FTIR based diagnosis would help decrease the cost through early intervention. FTIR based diagnosis of diseases, microbes, and healthy and unhealthy conditions is thus a future possibility that is less destructive and harmless compared to conventional methods (Toubas et al 2006). Similar approach can be used to diagnose microbial presence in blood samples (Maquelin et al 2003).

Fig. 6. Evaluation of polyposis in biopsies using the biomarkers derived from crypts. The Crypts from patients 1 and 2 had a polyposis while those of patients 3 and 4 did not. The values in the y axis are derived from the ratios of the integrated absorbance in the region 900-1185 cm\(^{-1}\) obtained at the middle and top of the crypt. The color codes are independent and used only to distinguish crypts from one another. Crypts with a value less than 1 are indicative of a propensity to have polyposis. At least 2 such abnormal crypts are required for assigning the worst prognosis.

Another clinically relevant application of FTIR spectroscopy related to the, in vivo diagnosis of diseases in human organs like colon, rectum, oesophagus, cervix, alimentary tract, nasopharynx and other areas deals with fiber optic based IR probes that can be
inserted to obtain spectra of the surface layers. Studies pertaining to these objectives have utilized fiber-optic sensors for diagnosis of malignancies (Li et al 2005, Katkuri et al 2010) and under conditions of aqueous interference (Bindig et al 2003). Presently malignant tissues are excised during colonoscopy and send for evaluation. With the development of suitable probes, it is envisioned that other than the area with the symptoms, adjoining areas can be evaluated to determine the extent of spread of disease and redefine the resection margin. Another potential use would be to differentiate diseases with similar symptoms like IBD and cancer using such probes. This would help in early intervention and prevent recurrent surgeries resulting in less patient discomfort and less expenses. Thus, the field of FTIR based disease diagnosis can be utilized in various levels such as the simple diagnosis using IR microscopes to advanced in-vivo methods using fiberoptic sensors.

9. Discussion

The diagnostic potential of FTIR spectroscopy relies upon its ability to detect and monitor changes in the biochemical fingerprint of abnormal cells, tissues and molecules as compared to the normal conditions along with its ability to monitor minute spectral differences among closely related samples. Thus, establishing reference set of spectra of every biological entity that is being studied is an integral part as with other spectroscopic methods, often the references data base built using model compounds, whose exact spectral contribution is known (Benedetti et al 1996). Biological samples are procured after many different types of procedures and thus, each type may demand different approach for spectral acquisition and analysis. In addition, the availability of the instrument set up and resources would play an important role in deciding which types of samples can be analyzed in a particular clinical set up. For example, when both FTIR microscopy and normal spectroscopy is available, blood components can be examined by the normal method while tissues and biopsies can utilize the microscopy. Similarly, the tissue architect dictates the type of measurement to be followed. While tissues like cervical epithelium where distinct zones permit the spectral acquisition at several locations within a uniform region (Figure 3) more complex tissues like melanoma and nevus make it imperative that the location is carefully selected to not exclude desired region and not include regions that do not meet the criterion of either normal or diseased conditions (Hammody et al 2005). Complicated tissues like colonic epithelium present two possibilities of data acquisition, along the cross section (Argov et al 2002,2004) or longitudinal section (Sahu et al 2004b, 2010) of the crypts and inevitably the availability of microscopic or FPA facilities. Understanding the spectral differences due to instrumentation has been undertaken with the objective of tailoring the techniques to suit set ups and also verifying the universal applicability of certain parameters compared to others (Hammody et al 2007).

The biogenesis of each cell or tissue needs to be studied before a suitable methodology is designed to obtain spectral data for the specific purpose it is intended to serve. The various protocols are also required to be suitably modified depending upon whether a sample is fresh, frozen or paraffin embedded to exclude interfering substances (Sahu et al 2005). The spectroscopy of microorganisms on the other hand may be simpler (Sahu et al 2006b). Both FTIR and Raman spectroscopy use vibrational spectroscopy and can be used independently or in combination with each other as a means of spectroscopic evaluation of biological samples (Mourant et al 2003b, Krishna et al 2005, Oust et al 2006). The prospective of
extending Fiberoptic detection systems to different organs is a future potential (Lucas et al 2005).

10. Conclusion
The technique of FTIR spectroscopy is now being used for different purposes like identification of organisms and their classification, determination of status of tissues, monitoring the effects of drugs on cell lines, monitoring treatment regimen and studying interactions between biological systems in the arena of biomedical research. Owing to its simplicity of data acquisition compared to more sophisticated methods like NMR and development of computational methods for rapid data processing, its relevance in biomedical research is bound to increase over the coming years, in spite of the initial block of having large amounts of spectra for analyses.

11. References


This informative and state-of-the-art book on Infrared Spectroscopy in Life Sciences designed for researchers, academics as well as for those working in industry, agriculture and in pharmaceutical companies features 20 chapters of applications of MIRS and NIRS in brain activity and clinical research. It shows excellent FT-IR spectra of breast tissues, atheromatous plaques, human bones and projects assessment of haemodynamic activation in the cerebral cortex, brain oxygenation studies and many interesting insights from a medical perspective.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following: