Study of cancer cell lines with Fourier transform infrared (FTIR)/vibrational absorption (VA) spectroscopy

Uceda Otero, E. P.; Eliel, G. S. N.; Fonseca, E. J. S.; Hickmann, J. M.; Rodarte, R.; Barreto, E.; Jalkanen, Karl J.

Published in:
Current Physical Chemistry

Publication date:
2013

Document Version
Early version, also known as pre-print

Link back to DTU Orbit

Citation (APA):
In this work we have used Fourier transform infrared (FTIR) / vibrational absorption (VA) spectroscopy to study two cancer cell lines: the Henrietta Lacks (HeLa) human cervix carcinoma and 5637 human bladder carcinoma cell lines. Our goal is to experimentally investigate biochemical changes and differences in these cell lines utilizing FTIR spectroscopy. We have used the chemometrical and statistical method principal component analysis (PCA) to investigate the spectral differences. We have been able to identify certain bands in the spectra which are so-called biomarkers for two types of cell lines, three groups for the 5637 human bladder carcinoma cell line (5637A, 5637B and 5637C), and another one for the HeLa human cervix carcinoma cell line. The vibrational modes can be assigned to specific bands involving characteristic motions of the protein backbone. This work shows that infrared vibrational absorption (VA) spectroscopy can be used as a useful tool in medical diagnostics that provides in principle additional information and detail to that which can be obtained/provided from conventional histological studies, and more conventional mass spectroscopic and NMR techniques. The use of high level vibrational spectroscopic simulations, in
addition to the chemometric and statistical tools of PCA, linear and quadratic discriminant analysis, and artificial networks methods that are good at finding correlations, but provide little if any physical, chemical and biochemical insight into the nature of the changes at a molecular level, is also strongly advocated and helpful to gain more physical, chemical and biological insight. Hence the combination of vibrational spectroscopic simulations and experimental vibrational absorption spectroscopy and imaging are advocated for future developments in this field.

Keywords

Principal component analysis (PCA), Quadratic discriminant analysis (QDA), Linear discriminant analysis (LDA), Fourier transform infrared (FTIR), Vibrational absorption (VA), Cervical cancer, Bladder cancer, Henrietta Lacks (HeLa) cancer cell line

1 Introduction

Cancer is one of the largest problems for public health authorities in the world. According to predictions based on current statistical data in the year 2030 the number of new cases of cancer will be of the order of 27 million and the number of deaths due to cancer will be approximately 17 million. About 60 percent of these new cases will occur in developing countries, even though it is known that about a third of these predicted new cases of cancer could be prevented, for example by immunizations and other preventive measures [1].

For the year 2012 in Brazil, 257,870 of the new cases of cancer in males and 260,640 in females are expected with the following breakdown by types: skin cancer no-melanoma (134 thousand), prostate (60 thousand), breast cancer in females (53 thousand), colorectal (30 thousand), lung (27 thousand), stomach (20 thousand), and uterine cervix in females (18 thousand) [2].

Being that cancer is the second leading cause of deaths worldwide [3], it is very important to be able to detect early and then to monitor the progression of various cancer types. The development of early stage detection and monitoring tools would allow the early detection of the early stages of cancer before invasive cancer stages/forms develop. In addition, these tools could be used to determine the types of treatment which are having the greatest effects during the various stages of the disease, that is, to monitor the diseased state as a function of radiation and/or chemotherapy, and other treatments with for example hormones or liposomes [4,5,6].

Vibrational absorption spectroscopy is a specific molecular spectroscopic tool that allows one to identify characteristic normal vibrational modes which are specific to specific molecular bonds, functional groups, and intra- and intermolecular
interactions (e.g., hydrogen bonding) as well as more complex molecular interactions (cation-π interactions, π-π interactions in DNA base stacking, dispersion forces and interactions, and through space versus through bond interactions). The main types of conventional (non-chiral sensitive) vibrational spectroscopy are Raman scattering and infrared absorption (IR), which give complementary spectral information due to the different selection rules. The chiral sensitive analogues are vibrational circular dichroism (VCD) [7] and Raman optical activity (ROA) [8]. Both chiral and nonchiral vibrational spectroscopies are powerful methods for investigating the chemical composition of samples, the former (VCD and ROA) able to be used to determine the absolute configuration of chiral molecules. In addition, the combination of both theoretical and experimental vibrational spectroscopic techniques has been used in to solve a number of biophysical and biochemical problems which either alone could not do [9-12]. Vibrational spectroscopy methods have been utilized in many studies for the detection of different types of cancer [13-25].

In the literature there are a number of diverse experimental vibrational spectroscopic investigations of cancer tissues that demonstrate the usefulness of these spectroscopic methods in understanding cancer progression from a molecular point of view which also document their potential application in cancer screening and diagnosis [26-31]

Raman spectroscopy has also been used to study cell lines derived from the cervix, revealing spectral variations mostly in peaks originating from DNA and proteins in cell lines expressing the E7 gene of HPV-16 compared to cells not affected by human Papillomavirus (HPV) [32].

In our work two cancer cell lines were studied: Henrietta Lacks (HeLa) human cervix carcinoma and 5637 human bladder carcinoma cell lines. The main objective of the experimental work was to investigate biochemical changes in these cancer cells line utilizing FT-IR spectroscopy and to see if one can distinguish between the various cancer cell lines using only the FTIR/VA spectra. In addition, we sought to identify and characterize so-called biomarkers (specific infrared bands) and to associate them with specific characteristic modes in either individual biomolecules, or biomolecular complexes, or aggregation states.

2 Experimental details

A total of 200 spectra were collected of two types of cell lines, and then put in groups of fifty spectra. The carbon dioxide peaks were removed from the spectra raw utilizing the commercial Origin program, version 8.0 and subsequently normalized for statistical analysis which was undertaken with the Minitab program, version 15.
2.1 Cell culturing

HeLa (human cervix carcinoma) cell and 5637 cell lines (5637A, 5637B and 5637C), human bladder carcinoma cell, were obtained from the Rio de Janeiro Cell Bank (PABCAM, Federal University of Rio de Janeiro, Rio de Janeiro, RJ, Brazil) cell culture collections. Cells were routinely maintained in Dulbecco’s modified Eagle's medium with 10% fetal calf serum, 1% L-glutamine and 1% penicillin-streptomycin (all from Sigma, St. Louis, MO, USA), under a humidified 5% CO2 atmosphere at 37°C. Cells were grown at a low density of 5x105 cells/35 mm culture dish (Nunc, Copenhagen, Denmark). HeLa and 5637 cells (After 12, 24 and 48 hours) were submitted to spectroscopic measurements.

2.2 Sample preparation

Prior to spectroscopic measurements the cells were detached using 0.25% trypsin and EDTA (1mM) for 3–5 min (all from Sigma, St. Louis, MO, USA) and centrifuged at 1200 rpm for 5 min. After removing the supernatant, cells were washed twice in PBS and centrifuged at 1200 rpm for 5 min after each wash. Finally, the supernatant was removed and cell pellets were plated on glass slides in 48 well plates for 24 h. The number of cells cultured to one slide was in the range of 1x106 cells. Before spectroscopic measurements, cells were washed twice in PBS and dried at room temperature to minimize contributions from water. All cell lines were manipulated exactly the same manner, allowing comparison of the spectroscopic features of the cell lines investigated.

2.3 FTIR measurements

Fourier Transform - Infrared (FTIR) measurements were performed using a FTIR spectrophotometer IR-Prestige-21, Shimadzu in Transmission Mode. Spectra were collected as an average of 64 scans with a resolution of 4.0 cm⁻¹. FTIR signals were accumulated over a spectral range of 700-4000 cm⁻¹. An aperture of 100x100 μm was set on groups of 2 types of cells. Four pellets, one of HeLa and three of bladder cell lines, were investigated.

2.4 Data analysis

Pre-processing of the raw FTIR spectra included the exclusion of the carbon dioxide peak were done with origin 8.0, and after the data analysis was performed in Minitab version 15. Principal component Analysis (PCA) was applied to the data set. PCA is a multivariate statistical technique used for to reduce the dimensionality and exploratory analysis the data. This technique has been used for spectroscopists allowing for objective interpretation of vibrational spectroscopic data [33].
Results and discussion

3.1 Statistical Analysis

The principal components (PC) were computed using a full range of FTIR spectra, between 1000 and 4000 cm\(^{-1}\), and a covariance matrix. Figure 1 shows the loading plot for the first three of the four PCs, which correspond to the variation of PCs as a function of wavenumber.

The PC1 represents 97.6% of the data variance. The main contribution of vibrational modes and their respective assignments are shown in the Table 1 [21,22]. The second PC had 2.0% of the data variance and the principal contributions were from the oligosaccharide C-OH stretch modes, the protein amide I mode, and the H-O-H deformation modes of water [24, 35-37]. PC3 with 0.3% of the data variance also showed changes in the C-C and C-O stretch modes (proteins), the amide I and II modes, and, in addition, contributions from the N-H symmetric stretch and the O-H asymmetric stretch vibrational modes [25-27]. Besides this, PC4 with 0.01% of the data variance showed contributions for almost all constituents in the range, but there is a non-significant contribution in the range between 1,000 and 1,500 cm\(^{-1}\). In conclusion, the major changes observed in the spectra were related to proteins, primarily the amides bands (amide I & II). All principal components show variations for the peaks assigned to proteins, primarily the amide bands. For each PC, the amide I and amide II band regions exhibit peaks corresponding to variations of the different conformational (secondary) structures of the proteins. In PC1 and PC2, changes are observed of the amide I mode at 1633 cm\(^{-1}\) and 1640 cm\(^{-1}\), respectively. In PC3, changes to a more disordered phase are observed in the structure of the vibrations in the amide I and amide II regions (1413 cm\(^{-1}\) -1695 cm\(^{-1}\)).

To classify the PC1, PC2, PC3 and PC4 data from the 5637A, 5637B, 5637C and HeLa groups, they were analyzed by the methods of discriminant analysis. Two forms of discriminant analysis exist: linear discriminant analysis (LDA) and quadratic discriminant analysis (QDA). In linear discriminant analysis, the spectra are classified into a group if the squared distance (also called the Mahalanobis distance) of observation to the group center (mean) is the minimum, under the assumption that the covariance matrices are equal for all groups. For quadratic discriminant analysis, there is no assumption that the groups have equal covariance matrices and an observation is also classified into the group that has the smallest squared distance. However, the squared distance does not simplify into a linear function, and for this reason the analyses in this work were done with QDA and the classification results are shown on Table 2. The summary of the classification table shows that quadratic discriminant analysis identified correctly 100% of the spectra.
Fig. 1 Loading plot of the PCs for two types of cell lines

Table 1 Peak assignments for FTIR spectra data recorded for cancer cell lines

<table>
<thead>
<tr>
<th>Peak Reference</th>
<th>Wavenumber (cm(^{-1}))</th>
<th>FTIR peak assignments</th>
<th>Ref. No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Fig.2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>1051</td>
<td>C-O-C stretching of DNA and RNA</td>
<td>(21)</td>
</tr>
<tr>
<td>II</td>
<td>1059</td>
<td>2-Methylmannoside</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Oligosaccharide C-OH stretching band</td>
<td>(24)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mannose &amp; mannose-6-phosphate</td>
<td></td>
</tr>
<tr>
<td>III</td>
<td>1121</td>
<td>Symmetric phosphodiester stretching band RNA</td>
<td>(25)</td>
</tr>
<tr>
<td>IV</td>
<td>1595</td>
<td>Ring C-C stretch of fenyl (2) &amp; Methylated nucleotides</td>
<td>(26)</td>
</tr>
<tr>
<td>V</td>
<td>1413 &amp; 1560</td>
<td>Stretching C-N, deformation N-H, Deformation C-H &amp; Ring base</td>
<td>(22)</td>
</tr>
<tr>
<td>VI</td>
<td>1630</td>
<td>Ring C-C stretch of phenyl, Amide I region  (26)</td>
<td></td>
</tr>
<tr>
<td>----</td>
<td>------</td>
<td>-----------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>VII</td>
<td>1633</td>
<td>C=C uracil, C=O  (22)</td>
<td></td>
</tr>
<tr>
<td>VIII</td>
<td>1640</td>
<td>Amide I band of protein and H-O-H deformation of water  (23)</td>
<td></td>
</tr>
<tr>
<td>IX</td>
<td>1694</td>
<td>A high frequency vibration of an Antiparallel β-sheet of amide I  (27)</td>
<td></td>
</tr>
<tr>
<td>X</td>
<td>2100</td>
<td>A combination of hindered rotation and O-H bending (water)  (26)</td>
<td></td>
</tr>
<tr>
<td>XI</td>
<td>2178</td>
<td>C=N vibrations  (22)</td>
<td></td>
</tr>
<tr>
<td>XII</td>
<td>3177</td>
<td>Stretching N-H symmetric  (22)</td>
<td></td>
</tr>
<tr>
<td>XIII</td>
<td>3184</td>
<td>Stretching N-H symmetric  (22)</td>
<td></td>
</tr>
<tr>
<td>XIV</td>
<td>3339</td>
<td>Stretching N-H asymmetric  (22)</td>
<td></td>
</tr>
<tr>
<td>XV</td>
<td>3435</td>
<td>Stretching O-H asymmetric  (22)</td>
<td></td>
</tr>
<tr>
<td>XVI</td>
<td>3439</td>
<td>Stretching O-H asymmetric  (22)</td>
<td></td>
</tr>
</tbody>
</table>

**Fig. 2** Average spectra for 4 groups of cells

Figure 2 shows the average spectra for 4 groups of cell lines, observed major similarity in cell lines 5637 (A, B and C), excepted at a lower wavenumber range. The FTIR data from the 5637 groups showed spectra likeness. It is correct because the groups are the same cell line. The peaks were sharp with the same intensity at low values of full width at half maximum (FWHM) and there is a relevant peak around 1,636 cm\(^{-1}\). The presence of ring C-C stretch of phenyl, C=C uracil, C=O, amide I band protein and H-O-H
were observed [17, 18, 34]. Nevertheless, all the spectra have a peak between in the range 1,050 - 1,090 cm\(^{-1}\), being the position of these the major difference between them. These peaks represent Phosphate stretching modes of DNA and RNA [11, 16]. In addition, all the spectra had broad bands and the peak around 3250 cm\(^{-1}\) with high intensity.

Figure 3 shows the tree-like diagram (dendrogram, using an average linkage and Euclidean distance of PC1, PC2, PC3 and PC4), which groups spectra into clusters according to similarity level. The sample 5637A is grouped in two subgroups, which also happens for the samples 5637B and 5637C. These samples were not different much, to the point that with a similarity about 80, clusters is similar. Nevertheless, in the HeLa sample we can see various subgroups with similarity smaller than 83, in total 7 sub-groups. Of these, 2 subgroups were more similar with the measures of the sample 5637 that with other subgroups of HeLa. We can see this similarity when we observed 3D-scatterplot; some measures of the HeLa sample have its main similar components to the sample 5637C. All groups were correctly separated, but they were divided additionally into various subgroups. This result suggests a variation in the chemical composition of cell lines in these groups, which can be expected from different cell groups in different environments.
The principal component (PC) scatterplot for FTIR spectral data results presented in Fig. 4 shows a distinctive separation between the different cell lines. The groups 5637 (A, B and C) are grouped together, confirming observed similarities between these types cells based on the mean spectra analysis. The HeLa group of cell lines are clearly separated and dispersed of the 5637 cell line group. This scatterplot is consistent with PCA of FTIR spectra data.

**Table 2** Summary of classification of QDA

<table>
<thead>
<tr>
<th>Classified into</th>
<th>True Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5637A</td>
</tr>
<tr>
<td>5637A</td>
<td>50</td>
</tr>
<tr>
<td>5637B</td>
<td>0</td>
</tr>
<tr>
<td>5637C</td>
<td>0</td>
</tr>
<tr>
<td>HeLa</td>
<td>0</td>
</tr>
<tr>
<td>Total N</td>
<td>50</td>
</tr>
<tr>
<td>N correct</td>
<td>50</td>
</tr>
<tr>
<td>Proportion</td>
<td>1.000</td>
</tr>
</tbody>
</table>
5. Conclusions

A promising classification of spectra FTIR was obtained using principal component analysis (PCA). The comparison between all FTIR spectra studied showed important differences at lower wavenumber which could be associated with the reduction and oxidation processes in the cancer cell lines. We can see the presence of phenylalanine and amide I bands assigned to the protein in all groups. Therefore, FTIR spectroscopy techniques have been and can be applied to the analysis of cancer cell lines delivering important information about the biochemical composition of the investigated samples. In addition, the use of high level first principles \textit{ab initio} wave function theory (AI-WFT) and Kohn Sham density functional theory (KS-DFT) in combination with vibrational, and NMR spectroscopy and incoherent inelastic neutron scattering (IINS) experiments have been shown recently to be useful in explaining the effects of various biological environments on both structures and species of biomolecules [9-12,38-42]. There more sophisticated theoretical methods can be used to interpret the changes seen in IR and Raman spectra of cancer and other diseased state tissues, cells and fluids versus those in the healthy state. PCA, LDA and other chemometric methods have been used to identify correlations and changes, but they have not been helpful in identifying the changes to individual residues and conformational and species (ionization and/or aggregation state) which would be necessary to design pharmaceutical agents which can be used in treatment regimes. Clearly vibrational spectroscopy and imaging are tools which can have a major impact in quantum molecular biology, biophysical chemistry, and molecular medicine. What has limited their use in medical applications is the complexity of the theoretical treatment and the nontransparent interpretation of the changes seen, observed and being used as so-called biomarkers of the diseased states. Here we advocate that the next generation of simulations needs to be at the quantum mechanical level using state of the art AI-WFT and KS-DFT methods which treat the effects due to the local environment quantum mechanically [12,43], and not classically [42], and also calculate the linear and nonlinear response properties which will give not only the structure and frequencies, but also the VA, VCD, Raman and ROA intensities and their changes, which contain a wealth of information concerning the interactions between the various biomolecules and their environment [11-13,44,45], changes that have been shown to lead to a number of diseases, cancer being one of the most prevalent and the one we have investigated here.

Acknowledgments

The authors would like to thank coordination for the Improvement of Higher Level – or Education Personnel (CAPES) of Brazil for their financial support on Projects Nanobiotec-661/2009 - CAPES. KJJ would like to thank the German Cancer Research Center (DKFZ) in Heidelberg, Germany and the Deutscher Akademischer Austausch Dient (DAAD) in Bonn, Germany for infrastructure, health insurance, and financial support which allowed Dr. Jalkanen to visit the Division of Functional Genome Analysis from September 2011 to January 2012, and to DTU Nanotech and the Technical University of Denmark for providing housing and living expenses during KJJ 3 month stay in the Department of Micro- and Nanotechnology. KJJ would also like to acknowledge Dr. Jörg Horshiel, Head of the Division of Functional Genome Analysis, and his group, especially Dr. Michaela Knapp-Mohammady, for discussions, seminars, and providing a wonderful and stimulating work environment.
References


[12] Jalkanen, K.J.; Degtyarenko, I.M.; Nieminen, R.M.; Nafie, L.A.; Cao, X.; Zhu, F.; Barron, L.D. Role of hydration in determining the structure and vibrational spectra of L-


