



Application of FT-IR spectroscopy on breast cancer serum analysis



Fatemeh Elmi^a, Afshin Fayyaz Movaghar^b, Maryam Mitra Elmi^{c,*},
Heshmatollah Alinezhad^d, Novin Nikbakhsh^e

^a Department of Marine Chemistry, Faculty of Marine & Oceanic Science, University of Mazandaran, Babolsar, Iran

^b Department of Statistics, Faculty of Mathematical Sciences, University of Mazandaran, Babolsar, Iran

^c Cellular and Molecular Biology Research Center, Health Research Institute, Babol University of Medical Sciences, Babol, Iran

^d Department of Chemistry, Faculty of Organic Chemistry, University of Mazandaran, Babolsar, Iran

^e Surgery Department, Faculty of Medicine, Babol University of Medical Sciences, Babol, Iran

ARTICLE INFO

Article history:

Received 29 October 2016

Received in revised form 14 June 2017

Accepted 15 June 2017

Available online 20 June 2017

Keywords:

FT-IR spectroscopy

Breast cancer

Serum

PCA-LDA analysis

ABSTRACT

Breast cancer is regarded as the most malignant tumor among women throughout the world. Therefore, early detection and proper diagnostic methods have been known to help save women's lives. Fourier Transform Infrared (FT-IR) spectroscopy, coupled with PCA-LDA analysis, is a new technique to investigate the characteristics of serum in breast cancer. In this study, 43 breast cancer and 43 healthy serum samples were collected, and the FT-IR spectra were recorded for each one. Then, PCA analysis and linear discriminant analysis (LDA) were used to analyze the spectral data. The results showed that there were differences between the spectra of the two groups. Discriminating wavenumbers were associated with several spectral differences over the 950–1200 cm^{-1} (sugar), 1190–1350 cm^{-1} (collagen), 1475–1710 cm^{-1} (protein), 1710–1760 cm^{-1} (ester), 2800–3000 cm^{-1} (stretching motions of $-\text{CH}_2$ & $-\text{CH}_3$), and 3090–3700 cm^{-1} (NH stretching) regions. PCA-LDA performance on serum IR could recognize changes between the control and the breast cancer cases. The diagnostic accuracy, sensitivity, and specificity of PCA-LDA analysis for 3000–3600 cm^{-1} (NH stretching) were found to be 83%, 84%, 74% for the control and 80%, 76%, 72% for the breast cancer cases, respectively. The results showed that the major spectral differences between the two groups were related to the differences in protein conformation in serum samples. It can be concluded that FT-IR spectroscopy, together with multivariate data analysis, is able to discriminate between breast cancer and healthy serum samples.

© 2017 Published by Elsevier B.V.

1. Introduction

Breast cancer is the most widespread malignant tumor in women around the world. The International Agency for Research on Cancer (IARC) reported that 1.7 million cases, 11.9% of the total diagnosed cancers worldwide, were those suffering from breast cancer, and that the incidence rates are on the rise in many parts of the world. A good illustration for that is Iran. It should be mentioned that the incidence of the breast cancer in Iran is increasing, and the age of getting afflicted to it is declining. Breast cancer is, very unfortunately, the most common cause of cancer death among Iranian women. The current screening methods of the breast cancer are the X-ray mammography and ultrasound assessments, but the gold standard is still the removal of a biopsy [1]. The emphasis on the early detection and the truthful diagnostic

methods has been known to help save women's lives. Breast cancer could be divided into benign and malignant lesions. The benign lesions are related to fibrocystic changes, whereas most malignant lesions are due to ductal carcinomas [2]. In order to the patho-physiological condition in human body influence the serum composition, the content of the serum, during sickness, should be thoroughly analyzed via such an innovative technique as FT-IR spectroscopy [3]. In recent years, FT-IR spectroscopy has been considered to be used as a tool for screening cells [4], tissues [5] and biological fluids [6,7]. FT-IR spectroscopy is a highly reproducible analytical technique for studying the structure and the conformation of such biomolecules as proteins and peptides [8], nucleic acids [9], carbohydrates, and lipids [10,11].

The concentrations of such serum components as total protein, cholesterol, glucose, and urea were measured using mid-IR spectroscopy [12,13]. Backhaus et al. used mid-IR spectroscopy for the detection of breast cancer [1]. Some other groups also reported the use of IR spectroscopy for the recognition of breast cancer [14–16]. Carcinoembryonic antigen (CEA), cancer antigen 15-3 (CA 15-3), the extracellular domain (HER-2), and tissue polypeptide antigen (TPA) are all regarded as tumor markers, but the sensitivity and the specificity of these markers are low,

* Corresponding author at: Cellular and Molecular Biology Research Center, Babol University of Medical Sciences, Babol, Iran.

E-mail addresses: f.elmi@umz.ac.ir (F. Elmi), a.fayyaz@umz.ac.ir (A.F. Movaghar), m.elmi@mubabol.ac.ir (M.M. Elmi), heshmat@umz.ac.ir (H. Alinezhad), novinsu@hotmail.com (N. Nikbakhsh).

especially for early stages of the disease [17]. In a study, using surface-enhanced Raman scattering of the breast cancer serum combined with multivariate data analysis, PCA-LDA was suggested as a technique to be used to distinguish the breast cancer from the healthy cases, and also to identify breast cancer at different stages [18]. The focus of this study was on the FT-IR spectra combined with multivariate data analysis, and the PCA-LDA for the analysis of human serum with the aim of detecting spectral parameters, which could be used as a non-invasive method for monitoring breast cancer in patients.

2. Materials and Methods

2.1. Samples and Sample Preparation

0.5 ml serum samples were collected from the sampled women with breast cancer prior to their surgery. Sera from blood donors, including healthy people who had come to be checked through mammography and sonography, were used as the control group. All participants were fully informed and given the consent forms prior to this study. The study was approved by the local ethics committee.

The mean age of the breast cancer patients was 48.88 ± 10.4 . The oldest patient was 73, and the youngest one was 20. The mean age of the blood donors was 44.84 ± 5.2 . The youngest was 34, and the oldest was 60. 43 sera were used from each group. All patients were housekeepers, except for four who were employees. They did not take any medication (only 6 took medication for about 10 years). They did not suffer from any diseases, nor did they have relatives with cancer problem (except for two cases). The stage of carcinomas was different from stage I to stage VI. The blood samples were centrifuged at 4°C , and then the serum samples were frozen at -80°C until the analysis stage. 5 μl of the serum sample was spread on 2 cm in diameter of the zinc selenide crystal. The sample was dried at the room temperature for about 10 min, and was then transferred to the FT-IR spectrometer.

2.2. FT-IR Spectral Data Processing

Infrared spectra were recorded with fourier transform infrared spectrometer (FT-IR, Bruker Tensor 27). All samples were measured in transmission mode, in the frequency range of $400\text{--}4000\text{ cm}^{-1}$ and the resolution of 4 cm^{-1} . The baseline correction and normalization were performed for all the spectra by OPUS software. The baseline correction was done through rubber band method, and the normalization was performed by vector method. For normalization, the data were transformed to unit variance through being divided by the standard deviation. The data analysis was performed using R-3.4.0 software. From the original spectra of the cancerous samples and those of the controls, a mean spectrum was formed and the standard deviation was calculated. All spectra, outside the mean value of ± 2 standard deviations, were excluded from the analysis.

To build a classification method, linear discriminate analysis (LDA) was combined with PCA (principal component analysis), which is called PCA-LDA. At first, with the use of PCA method, the number of variables from a large number was reduced to a small number of variables (which were the main components), then with the use of LDA method, a suitable model for grouping data based on independent variables (wavelengths) was developed. Five PCs, explaining 99.8% of the spectral variance, were selected. Then, LDA procedure was performed based on the selected PCs in various spectral regions.

3. Results and Discussion

A comparison of the averaged spectra between the pathological and control serum samples is shown in Fig. 1. The preliminary data demonstrated that the pathological serum samples were spectrally unique when they were compared with the control serum samples, especially in lipids and proteins regions. Nevertheless, it should be noted that

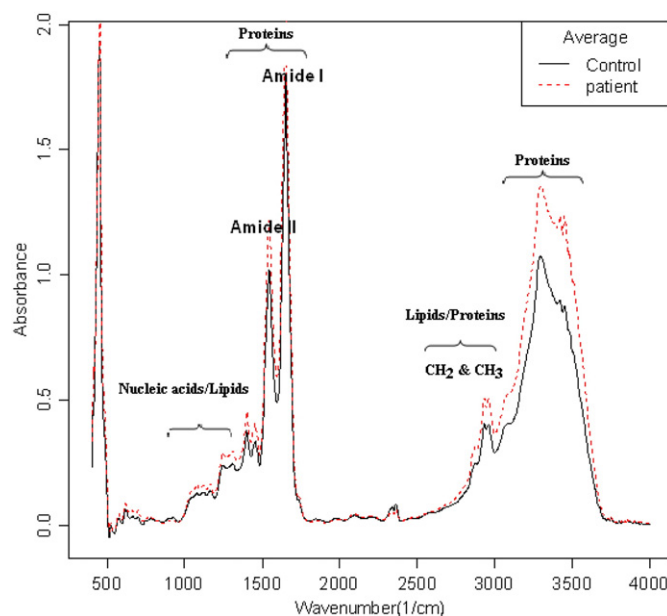


Fig. 1. The averaged FT-IR spectra of control and patients in the range of $4000\text{--}400\text{ cm}^{-1}$.

the correct spectral recognition of the dissimilarity requires the use of multivariate analytical methods. Several spectral ranges were analyzed with PCA-LDA, which are as follows: $1475\text{--}1710\text{ cm}^{-1}$ associated with proteins; $3000\text{--}3060\text{ cm}^{-1}$ associated with unsaturated lipids; $1710\text{--}1760\text{ cm}^{-1}$ associated with esters; $950\text{--}1200\text{ cm}^{-1}$ associated with sugar, $1190\text{--}1350\text{ cm}^{-1}$ associated with collagen; $1220\text{--}1250\text{ cm}^{-1}$ associated with phosphate diesters, $3090\text{--}3700\text{ cm}^{-1}$ associated with NH stretching; and $2800\text{--}3000\text{ cm}^{-1}$ associated with stretching motions of $-\text{CH}_2$ and $-\text{CH}_3$.

A separation between healthy and cancerous groups on the basis of serum samples spectra was carried out through PCA-LDA analysis, as demonstrated in Table 1 and Figs. 2–4 (Figs. S1–S5). Loadings plot can identify the segregating wavenumbers (as weightings). The loadings plot (Fig. 3a) derived from PCA-LDA of $3090\text{--}3700\text{ cm}^{-1}$ spectral range distinguished the four primary wavenumbers, which were important for the separation of two groups. These include 3297, 3424, 3435, and 3451 cm^{-1} . It is the region of the N—H stretching vibration assigned to

Table 1

Diagnostic accuracy, sensitivity, and specificity of PCA-LDA analysis results for the classification between serum from breast cancer patients and controls.

Sample	Wave length (cm^{-1})	Accuracy %	Sensitivity %	Specificity %
Control	3090- 3700	83	84	74
Patient		80	76	72
Control	1475- 1710	78	85	81
Patient		72	68	60
Control	1600- 1700	71	76	69
Patient		63	68	58
Control	1500- 1560	80	90	86
Patient		75	71	63
Control	2800- 3000	75	82	74
Patient		64	65	54
Control	3000- 3060	80	83	75
Patient		74	81	75
Control	1710- 1760	83	90	84
Patient		78	74	71
Control	950- 1200	77	81	74
Patient		69	74	63
Control	1190- 1350	78	86	77
Patient		68	68	57
Control	1250- 1220	82	85	76
Patient		73	78	71

Download English Version:

<https://daneshyari.com/en/article/5139555>

Download Persian Version:

<https://daneshyari.com/article/5139555>

[Daneshyari.com](https://daneshyari.com)