J. Med. Biol. Eng. (2016) 36:369–378
DOI 10.1007/s40846-016-0133-0

ORIGINAL ARTICLE

# Molecular-Level Characterization of Normal, Benign, and Malignant Breast Tissues Using FTIR Spectroscopy

**Wael M. Elshemey[1] · Alaa M. Ismail[1] · Nihal S. Elbialy[1,2]**

Received: 13 July 2015 / Accepted: 10 December 2015 / Published online: 7 June 2016


**Abstract** The present study investigates the diagnostic capability (sensitivity, specificity, and diagnostic accuracy) of the Fourier transform-infrared (FTIR) technique as a tool for the characterization of excised breast tissues with special emphasis on the differentiation between malignant and benign tumors. For such purpose, normal, malignant, and benign excised breast tissue samples were collected. The scatter diagrams of suggested FTIR characterization parameters were plotted and the corresponding diagnostic indices were calculated, showing maximum values reaching 100 % for differentiating between normal and diseased samples. The receiver operating characteristic analysis of such indices reported that the maximum area under the curve reached a value of one, which indicates highly accurate performance. The capability of the technique to differentiate between benign and malignant samples yielded promising results. A deconvolution of the Amide I protein peak revealed a high percentage of ordered protein secondary structure elements in normal tissue compared to those for malignant and benign tumors. The results indicate that the FTIR technique is ready for application in the characterization of excised breast tissues.

**Keywords** Breast cancer · Fourier transform-infrared (FTIR) spectroscopy · Receiver operating characteristic (ROC) curve · Scatter diagram · Deconvolution

✉ Wael M. Elshemey
welshemey@sci.cu.edu.eg

[1] Biophysics Department, Faculty of Science, Cairo University, Cairo University Rd., Giza 12613, Egypt

[2] Department of Physics, King Abdulaziz University, Jeddah 21589, Kingdom of Saudi Arabia

## 1 Introduction

Breast cancer is a heterogeneous disease that mostly affects 40–55 years old women [1]. The GLOBOCAN 2008 project of the World Health Organization estimated that breast cancer is the leading cause of cancer death among females, accounting for 23 % (1,380,000) of the total cancer cases and 14 % (458,400) of the total cancer deaths in 2008 [2]. It was reported that half of the breast cancer cases and 60 % of deaths occurred in economically developing countries. The international variation in incidence rates for breast cancer may be due to differences in reproductive and hormonal factors and the possibility of early diagnosis facilities [2]. The survival rate can be improved by performing accurate screening for breast cancer in its early stages.

The main motivation of this work arises from the fact that although there are several methods that are currently used for the diagnosis of breast cancer, such as histopathological examination, mammography, ultrasound imaging, and magnetic resonance imaging (MRI), they all have inaccuracies.

For example, mammography is inaccurate in detecting 10–14 % of clinically diagnosable cancers. It is less accurate in detecting breast cancer without micro calcification in compared to breast cancer with microcalcification women with high mammographic density [3–5].

Ultrasound imaging is less sensitive than mammography and has high false positive rates. Additionally, it is unable to image most ductal carcinoma in situ cases [6]. MRI is an expensive technique. Although it is sensitive for visualization of intra-ductal carcinoma and infiltrating ductal carcinoma when it is clinically and mammographically obscure, it is less specific in differentiating between benign

and malignant lesions [7]. The American Cancer Society (ACS), American College of Radiology (ACR), and Society of Breast Imaging (SBI) guidelines concur that screening with MRI is inappropriate for women with >15 % life-time risk for breast cancer [6].

Histopathological investigation is time-consuming, less precise for early stage malignancy detection, compared to late stage malignancy detection and may be subject to personal errors, which could have fatal consequences if used to select treatment for patients. Additionally, it cannot clarify the biochemical changes that take place during carcinogenesis [8–10].

In recent years, several medical and biological studies have been performed using spectroscopic techniques. Since vibrational spectra provide valuable information on the molecular scale, including structure and composition of biological materials, several applications in the medical field have been reported.

Infrared (IR) spectroscopy is considered a powerful tool in the study of various human cancers and the biochemistry of cancer [11–17]. Fourier transform-infrared (FTIR) spectroscopy is used to detect biomolecular changes in tissues based on the absorption of IR radiation arising from chemical bonds in the functional groups of molecules. Such frequency changes in the IR absorption spectrum correspond to the molecular structure, molecular conformation, and intermolecular interactions. The peak frequencies are used to monitor changes in chemical structure. The availability of a certain structure is detected via relative peak intensities. FTIR spectroscopic analyses are currently focused on the histological characterization of breast carcinoma tissues [1, 18–20]. Advantages of the FTIR technique include simplicity, sensitivity, rapidity, reproducibility, non-destructiveness, low cost, minute sample amount requirement, and no staining or reagent requirements [1, 18, 19, 21]. FTIR is sensitive to the secondary structure of proteins. The analysis of IR spectra provides an accurate estimation of the content of protein secondary structures and gives information on the stability and dynamics of a protein structure [22–24].

Previous studies revealed that FTIR spectroscopy can differentiate between the grades of lymphoma, breast ductal carcinoma in situ, and invasive ductal carcinoma grades [21, 25]. Another study showed that the differences in IR absorption of basement membranes from tumor blood capillaries can aid in have an important role in tumor grading [26]. FTIR spectroscopy can distinguish the structural changes in prostate DNA associated with both aging and cancer development [27].

## 2 Materials and Methods

### 2.1 Sample Preparation and FTIR Measurement

Sixty-nine excised breast tissue specimens were collected over a period of about 9 months from women undergoing mastectomy. These samples were taken just before sample disposal. The samples are completely anonymous and cannot be related to specific patients. A histopathologist classified the specimens into 20 normal, 28 malignant, and 21 benign tissues. The samples were kept at −80 °C before they were lyophilized using a freeze dryer (Edwards freeze drier, England). Sample lyophilization is an essential step as water strongly absorbs IR radiation, masking protein peaks [28–30]. Lyophilized breast tissue samples were measured using an FTIR Nicolet 6700 spectrometer (Thermo Electron, Madison, WI, USA). The spectrometer made 16 scans per spectrum at a resolution of 2 $cm^{-1}$, covering the spectral range of 400–4000 $cm^{-1}$.

### 2.2 Data Processing

After area normalization of all taken spectra, the average FTIR spectrum for normal, malignant, and benign samples was calculated, respectively. Eight absorption peaks were selected at wavenumbers of around 1099, 1163, 1237, 1536, 1657, 2854, 2925, and 3456 $cm^{-1}$, respectively. These peaks were utilized to differentiate between normal, malignant, and benign samples. A group of suggested FTIR characterization parameters (peak center, peak intensity, area under peak, relative intensity, and relative area) were used to obtain quantitative values for the main features of various FTIR spectra. The peak centers and peak intensities were determined from the spectral data. The average area under the peak was calculated for each of the selected peaks using the trapezoidal integration method, calculated using Microsoft Excel 2010. The relative intensity and relative area were respectively calculated by dividing the band area and band intensity of the Amide I peak at 1657 $cm^{-1}$ (sensitive to protein secondary structure) by the band area and band intensity of each of the remaining seven selected peaks.

The Duncan multiple range test was performed using the IBM SPSS 17 statistics software package to determine whether the selected eight peaks could significantly differentiate between the three sample types using the five characterization parameters. These characterization parameters were considered as diagnostic test results characterization parameters were considered as diagnostic tests for the detection of each breast tissue type.

### 2.3 Diagnostic Indices of Scatter Diagrams

In order to calculate the sensitivity, specificity, and diagnostic accuracy of the diagnostic tests, scatter diagrams were plotted [31]. These diagrams show each breast tissue type on the X-axis and the corresponding individual values of a certain characterization parameter (diagnostic test) plotted on the Y-axis. A horizontal cut-off line was drawn twice. A line was drawn in order to differentiate between normal and diseased samples (mean ± 2SD of normal samples [23, 32, 33]; SD: standard deviation) and another line was plotted in order to differentiate between malignant and benign samples (mean of malignant samples [34]). The sensitivity of the diagnostic FTIR technique is the percentage of samples with the disease identified correctly by the test as positive (true positive divided by the sum of true positive and false negative).The specificity of the diagnostic test is the percentage of samples without disease specified correctly by the test as negative (true negative divided by the sum of true negative and false positive). The diagnostic accuracy is the percentage of samples correctly specified as true positive and true negative out of all tested samples (the sum of true positive and true negative divided by the total sample number) [31, 35–37].

### 2.4 Receiver Operating Characteristic Curves

In the present work, the receiver operating characteristic (ROC) curve was utilized as a standard method to investigate the performance of the diagnostic tests [34, 38]. The ROC curve is a plot of sensitivity versus 1-specificity. The curve was obtained by changing the cut-off values above and below the optimal cut-off and then calculating the corresponding sensitivity and specificity [38–41]. The area under the ROC curve (AUC) determines the preferable diagnostic characterization parameter. If AUC is less than 0.5, then the test is non-informative. If the value of AUC is in the range of 0.5–0.7, then the test is poorly accurate. A moderately accurate test has a value of 0.7–0.9 and a highly accurate test has a value of 0.9–1.0. A perfect test has an AUC value equal to unity [34, 40].

### 2.5 Second Derivative and Deconvolution of Amid I Band

After smoothing the area-normalized FTIR spectra with a seven-point Savitzky–Golay filter [42–45], the second derivative spectra were obtained using Origin 8.0 software. Two narrowing procedures (second derivative and deconvolution) were used to determine protein secondary structure components from the analysis of the Amide I band (1600–1700 $cm^{-1}$) [29, 30]. Since this band arises from the C=O stretching vibration of the peptide group and its frequency depends on hydrogen bonding and coupling of the polypeptide backbone, it is sensitive to protein conformation [46]. The spectral deconvolution of the Amide I absorption band was performed via the curve fitting method using Origin 8.0 software to determine the elements of protein secondary structure using the frequencies predicted by the second derivative [42, 47]. In the curve fitting method, the peak positions, widths, and heights corresponding to different protein secondary structure components were approximated using Gaussian functions [48].

## 3 Results

### 3.1 Averaged FTIR Spectra of Various Samples

Figure 1 shows the averaged FTIR spectra of normal, malignant, and benign breast tissue samples. The spectrum of normal tissues shows sharp absorption peaks at wavenumbers of 710, 1099, 1163, 1237, 1450, 1745, 2854, and 2925 $cm^{-1}$. The same peaks, but with lower intensities, are observed in the average spectra of the malignant and benign samples. The absorption peaks at wavenumbers of 1536, 1657, and 3456 $cm^{-1}$ have much lower intensities in the normal spectrum compared to those in the malignant and benign spectra. The assignment of main absorption peaks is presented in Table 1.

### 3.2 Characterization Parameters and Scatter Diagrams

Table 2 summarizes the characterization parameters and peaks that significantly differentiate the investigated breast tissue types. This table may be a guide for the use of the FTIR technique in the characterization of breast tissue samples. The availability of five useful successful

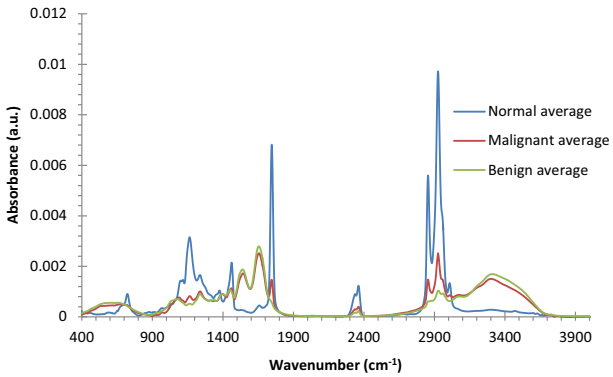


**Fig. 1** Average area normalized FTIR spectra of normal (*blue*), malignant (*red*), and benign (*green*) lyophilized human breast tissues

**Table 1** Peak assignments of FTIR spectra of normal, malignant, and benign human breast tissues

| | Normal | Malignant | Benign |
|---|---|---|---|
| Peak center | 1099.9 $cm^{-1}$ | 1086 $cm^{-1}$ | 1080.2 $cm^{-1}$ |
| Peak assignment | Symmetric $PO_2$ stretching (phosphate II) [50] | Symmetric phosphate stretching modes or stretching modes originating from the phosphor diester groups in nucleic acids, suggesting an increase in the nucleic acids in the malignant tissues [51]<br>V ($PO_2$ symmetric stretching of phosphodiesters) [52] | Symmetric phosphate ($PO_{2sym.}$) stretching [53]<br>Collagen and phosphor-diester groups of nucleic acids [54] |
| Peak center | 1163.4 $cm^{-1}$ | 1166.2 $cm^{-1}$ | 1163.1 $cm^{-1}$ |
| Peak assignment | C-O stretching band of collagen (type I) [55] | Mainly from the C-O stretching mode of C–OH groups of serine, threonine, and tyrosine of proteins [56] | C-O stretching band of collagen (type I) [55] |
| Peak center | 1237.6 $cm^{-1}$ | 1238.2 $cm^{-1}$ | 1238.3 $cm^{-1}$ |
| Peak assignment | Asymmetric $PO_2$ stretching (phosphate I) [50] | Asymmetric $PO_2$ stretching (phosphate I) [50]<br>Amide III [57] | Asymmeteric $PO_2$ stretching (phosphate I) [50] |
| Peak center | 1536 $cm^{-1}$ | 1543.1 $cm^{-1}$ | 1539.4 $cm^{-1}$ |
| Peak assignment | Stretching C = N, C = C [50] | Amide II absorption (primarily N–H bending coupled to a C-N stretching vibrational mode) [58] | |
| Peak center | 1657.3 $cm^{-1}$ | 1655.4 $cm^{-1}$ | 1654.5 $cm^{-1}$ |
| Peak assignment | Alpha-helical structure of Amide I [9]<br>Amide I absorption (predominantly C = O stretching vibration) | | |
| Peak center | 1745.8 $cm^{-1}$ | 1744.1 $cm^{-1}$ | 1743.3 $cm^{-1}$ |
| Peak assignment | Ester group (C = O) vibration of triglycerides [59] | | |
| Peak center | 2854.1 $cm^{-1}$ | 2855.7 $cm^{-1}$ | 2855.6 $cm^{-1}$ |
| Peak assignment | Symmetric stretching vibration of $CH_2$of lipids [60] | | |
| Peak center | 2925.4 $cm^{-1}$ | 2926.1 $cm^{-1}$ | 2928.5 $cm^{-1}$ |
| Peak assignment | C-H stretching bands in normal tissues [59]<br>Stretching C-H [50]<br>Asymmetric stretching vibration of $CH_2$ of lipids | | |
| Peak center | 3456.5 $cm^{-1}$ | 3301.3 $cm^{-1}$ | 3315.3 $cm^{-1}$ |
| Peak assignment | N–H stretching bands<br>Amide A band [9] | | |

parameters for breast tissue characterization necessitated the application of additional tests for further judging and comparing the diagnostic capabilities of each parameter to determine the best one. For such purpose, the two peaks (1657 $cm^{-1}$: Amide I and 2925 $cm^{-1}$: asymmetric stretching vibration of $CH_2$ of lipids) common among all of the characterization parameters (except peak center) were selected for analyzing their performance (peak intensity and are for 1657 and 2925 $cm^{-1}$ peaks and 1657/2925 $cm^{-1}$ ratio) using scatter diagrams. Figure 2 shows an example of a scatter diagram of two characterization parameters (peak intensity and area for the, of 2925 $cm^{-1}$ peak). The solid horizontal line is the optimal cut-off for differentiating between normal and diseased breast tissue samples. The dotted horizontal line represents the optimal cut-off for differentiating between malignant and benign breast tissue samples.

### 3.3 Diagnostic Indices

The quantitative evaluation of the performance of each characterization parameter and its corresponding optimal cut-off each couple of a characterization parameter and an optimal cut-off can be achieved using the diagnostic indices sensitivity, specificity, and diagnostic accuracy. The values of diagnostic indices for the characterization of normal and diseased (malignant and benign) samples show an extremely high sensitivity for all characterization parameters at all peaks, with peak ratios reaching a value of 100 %and having a minimum of 97.9 % (intensity of 2925 $cm^{-1}$ peak). Moreover, the specificity for the same test is also very high, reaching a maximum of 100 % (for area under 2925 $cm^{-1}$ peak) and having a minimum of 90 % (peak intensity and area of 1657/2925 $cm^{-1}$ and peak intensity of 1657 $cm^{-1}$). The diagnostic accuracy provides

**Table 2** Characterization parameters and peaks that significantly ($p < 0.05$) differentiate between investigated breast tissue types

| Characterization parameter | Differentiates three breast tissue samples | Differentiates normal from diseased (malignant and benign) only | Differentiates malignant from benign only |
|---|---|---|---|
| Peak center | 1099 $cm^{-1}$ | 1237 $cm^{-1}$ | 1163 $cm^{-1}$ |
| | 1536 $cm^{-1}$ | 1657 $cm^{-1}$ | 2925 $cm^{-1}$ |
| | | 3456 $cm^{-1}$ | |
| Peak intensity | 1163 $cm^{-1}$ | 1099 $cm^{-1}$ | |
| | 1657 $cm^{-1}$ | 1237 $cm^{-1}$ | |
| | 2854 $cm^{-1}$ | 1536 $cm^{-1}$ | |
| | 2925 $cm^{-1}$ | 3456 $cm^{-1}$ | |
| Area under peak | 1657 $cm^{-1}$ | 1099 $cm^{-1}$ | |
| | 2854 $cm^{-1}$ | 1163 $cm^{-1}$ | |
| | 2925 $cm^{-1}$ | 1237 $cm^{-1}$ | |
| | | 1536 $cm^{-1}$ | |
| | | 3456 $cm^{-1}$ | |
| Relative peak intensity | 1657/1099 | 1657/1536 | |
| | 1657/1163 | 1657/3456 | |
| | 1657/1237 | | |
| | 1657/2854 | | |
| | 1657/2925 | | |
| Relative area under peak | 1657/1099 | 1657/3456 | |
| | 1657/1163 | | |
| | 1657/1237 | | |
| | 1657/2854 | | |
| | 1657/2925 | | |

insight into the overall performance. The diagnostic accuracy for the differentiation of normal and diseased breast tissues using the present technique reaches a maximum of 100 % (area under 2925 $cm^{-1}$ peak) and a minimum of 97.1 %(peak intensity and area of 1657/2925 $cm^{-1}$ and peak intensity of 1657 $cm^{-1}$).

The diagnostic indices for differentiating between malignant and benign tumors show some very promising sensitivity results. A maximum sensitivity of 100 % was obtained for four parameters (intensity and area of 2925 $cm^{-1}$ peak and intensity and area of 1657/2925 $cm^{-1}$ ratio) and minimum sensitivities of 76.1 and 85.7 % were obtained for the intensity and area of the 1657 $cm^{-1}$ peak, respectively. Unfortunately, the calculated specificities are very low, reaching a maximum of only 50 % (intensity of 1657/2925 $cm^{-1}$ ratio) and a minimum of 25 % (area of 2925 $cm^{-1}$ peak). The overall diagnostic accuracy parameter is not too low. It reaches a maximum of 71.4 % (intensity of 1657/2925 $cm^{-1}$ ratio) and a minimum of 53.1 % (area of 2925 $cm^{-1}$ peak).

### 3.4 Receiver Operating Characteristic Curves

The accuracy of the previous results can be evaluated using the ROC curves. The AUC is the parameter of interest. For the characterization parameter used to differentiate between normal and diseased breast tissue samples, the AUC for all parameters yielded a value of 1. This AUC value indicates that the test is perfect in differentiating between normal and diseased tissues according Greiner et al. [40] and Elshemey et al. [34].

Figure 3 shows an example of the ROC curve for the peak intensity and area of 2925 $cm^{-1}$ for differentiating between malignant and benign breast tissue samples. The AUC values of the intensity and area of the 2925 $cm^{-1}$ peak and 1655/2925 $cm^{-1}$ ratio are 0.927, 0.913, 0.898, and 0.908, respectively, indicating that these parameters are highly accurate in differentiating between malignant and benign breast tissue samples [34, 40]. The intensity and area of the 1657 $cm^{-1}$ peak show AUC values of 0.670 and 0.726, respectively, indicating that these two parameters are poorly and moderately accurate, respectively. The ROC curves demonstrate the superior performance and accuracy of the present technique in differentiating between normal and diseased breast tissue samples. It also shows that the present technique is highly or moderately accurate in differentiating between malignant and benign breast tissue samples, a result which supports the usefulness of the present technique in differentiating between diseased breast tissue samples.

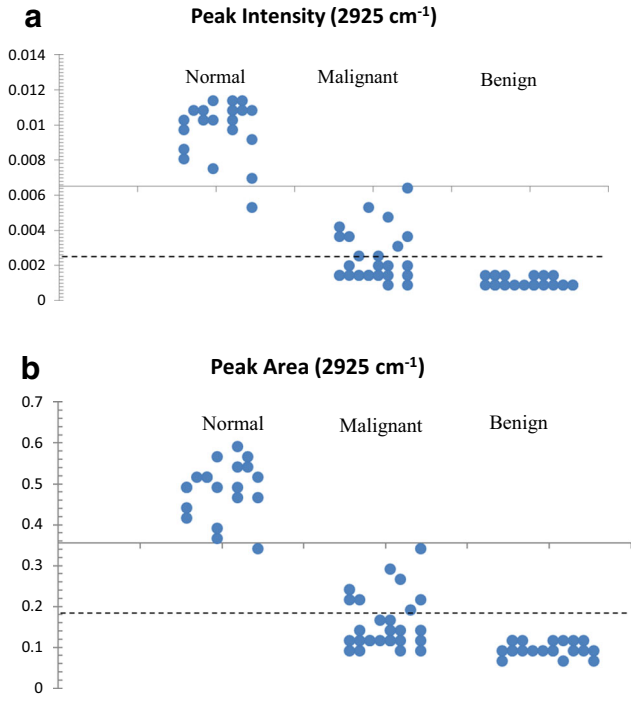


**Fig. 2** Scatter diagrams of peak intensity and peak area parameters, respectively, at 2925 $cm^{-1}$ absorption peak. *Solid horizontal line* is optimal cut-off line for differentiating between normal and diseased (malignant + benign) breast tissue samples. *Dotted horizontal line* is optimal cut-off line for differentiating between malignant and benign breast tissue samples

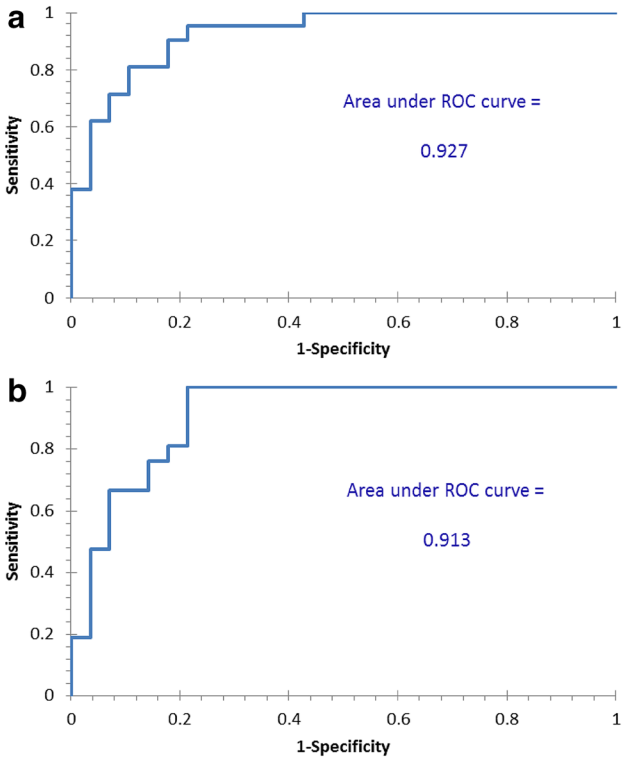


**Fig. 3** ROC curves of peak intensity and peak area parameters, respectively, at 2925 $cm^{-1}$ absorption peak for differentiating between malignant and benign breast tissues

### 3.5 Second Derivative Spectrum and Deconvolution of Protein Amide I Peak

Important information on the secondary structure of proteins (α-helix, ß-sheet, ß-turns, random coils, and aggregated strands) can be obtained by calculating the second derivative spectrum of the protein Amide I peak range (1600–1700 $cm^{-1}$) and then deconvoluting the Amide I peak via multi-Gaussian peak fitting.

The second derivative is a band narrowing technique that helps to determine the peak frequencies constituting the Amide I peak [42, 47]. These frequencies are then used in the deconvolution process, where the convolved Amide I peaks are totally resolved and information about peak centers, areas, widths, and heights becomes available. The second derivative spectra and peak frequency values for benign breast tissue are shown in Fig. 4a. The corresponding deconvoluted peaks are given in Fig. 4b.

The percentage of secondary structure content of the deconvolved peaks for the three breast tissue types are shown in Table 3. Several differences exist between normal and diseased (malignant and benign) samples. The $3_{10}$ helix and aggregated strands are absent in normal tissue. Normal tissue has higher α-helix and ß-sheet content (40.3 and 57.3 %, respectively) compared to that of malignant (5.8 and 22.5 %, respectively) and benign (7.3 and 28.6 %, respectively) samples. The percentages of unordered structures in diseased tissues are 48.0 % for benign and 26.7 % for malignant breast tissue samples compared to the normal tissue value of only 2.4 %. There are some differences between malignant and benign breast tissues that can be seen in Table 3. Benign tissue is characterized by a higher content of unordered coils, a small percentage of aggregated strands, and the absence of beta sheets.

## 4 Discussion

### 4.1 Differentiation of Breast Tissues Using Their Average FTIR Spectra

The large differences between the average spectra of normal tissue and malignant and benign tissues (Fig. 1) provide strong support for the possibility of differentiating between normal and diseased tissues using a simple comparison of their FTIR spectra. The suggested FTIR characterization parameters (Table 2) show superior capability in differentiating between different breast tissue samples. The first column in Table 2 shows that the characterization

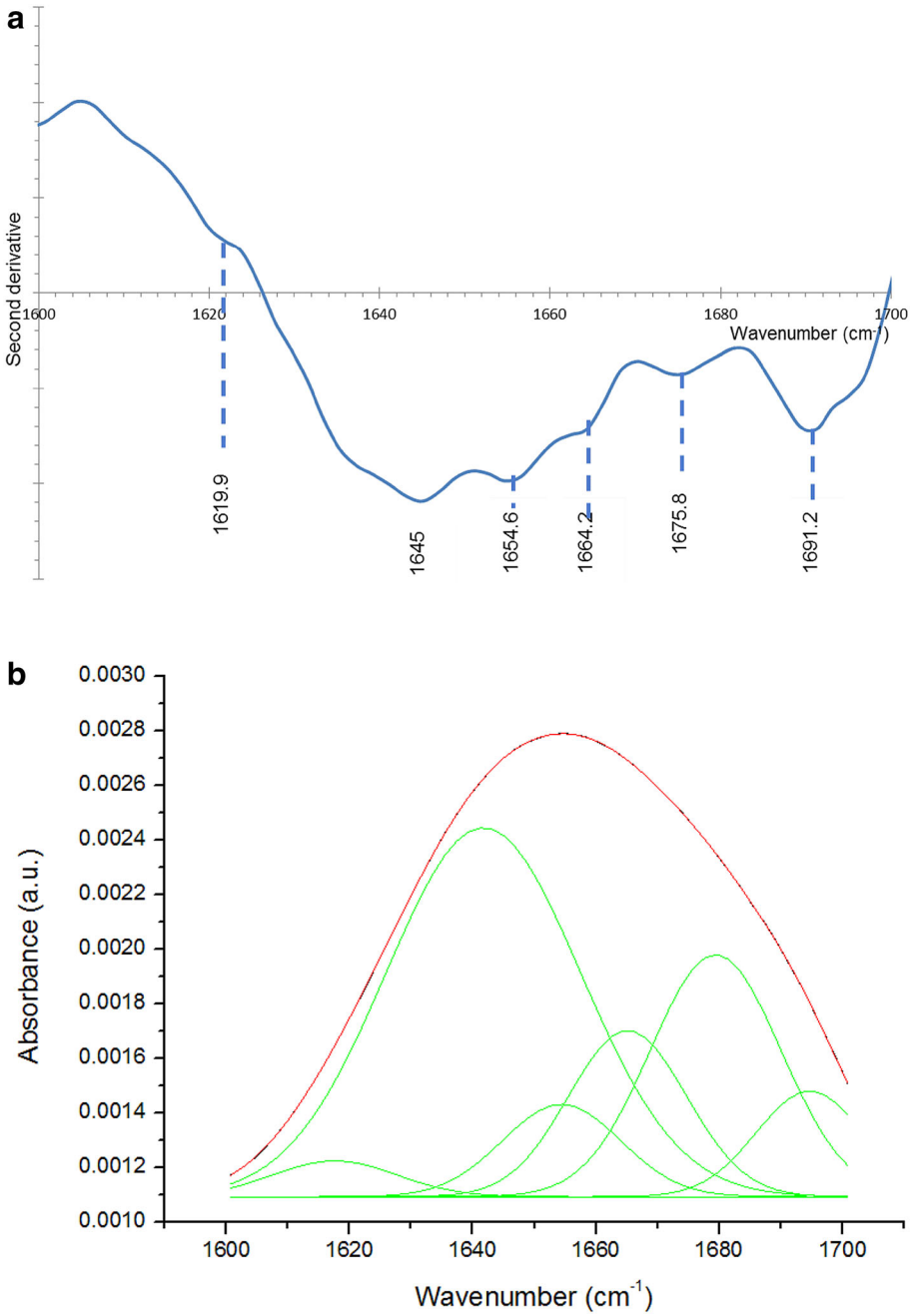


**Fig. 4** **a** Second derivative of averaged FTIR spectra for benign lyophilized human breast tissues in frequency region of 1600–1700 $cm^{-1}$. **b** Curve fitting into six-Gaussian-peak components deconvolved components (*green curves*), measured absorption spectra (*black*), and curve-fitted spectrum (*red*) (reduced $\chi^2$. 1.5 E-12, correlation $R^2 = 0.999$)

**Table 3** Percentage of protein secondary structures for Amide I (1600–1700 $cm^{-1}$) absorption band derived from deconvolved peaks

| Secondary structure (28) | Normal (%) | Malignant (%) | Benign (%) |
|---|---|---|---|
| Alpha helix (1648–1660 $cm^{-1}$) | 40.3 | 5.8 | 7.3 |
| $3_{10}$ helix (1660–1670) | | 25.9 | 13.1 |
| Beta sheet and beta turns (1625–1640 $cm^{-1}$) and (1670–1695 $cm^{-1}$) | 57.3 | 22.55 | 28.6 |
| Unordered structures (1640-1648) | 2.37 | 26.7 | 48.03 |
| Aggregated strands (1610–1628 $cm^{-1}$) | | | 2.97 |

parameters (peaks and peak ratios) are capable of differentiating between the three tissue types.

### 4.2 Differentiation of Breast Tissues Using Optimal Cut-Offs of Scatter Diagrams

The scatter diagrams (for example, Fig. 2) show that the optimal cut-off between normal and diseased samples is always better (showing large numbers of true positive and true negative samples and low numbers of false positive and false negative samples) than the optimal cut-off between malignant and benign samples. This observation reflects lower capability in differentiating between malignant and benign samples compared to differentiating between normal and diseased samples.

### 4.3 Differentiation of Breast Tissues

The extremely high sensitivity of all characterization parameters at all peaks and peak ratios of the diagnostic indices for the characterization of normal and diseased (malignant and benign) samples strongly supports the application of this technique in conjunction with conventional pathology for the diagnosis of excised breast tissue abnormalities. The obtained values of sensitivity and specificity are potentially high compared to those reported for well known tumor markers are potentially high, knowing that, for example, the sensitivity and specificity for the golden tumor marker (Alpha-feto protein) of hepatocellular carcinoma does not exceed 70 % [35] (35). The area under the 2925 $cm^{-1}$ peak is probably the best characterization parameter studied in this work. This parameter emerging biomarker has shown 100 % sensitivity, specificity, and diagnostic accuracy.

Knowing that differentiating between malignant and benign tumors is a difficult issue [49, 50], the results of the diagnostic indices for differentiating between these two tissues using the present technique provide supportive information for a histopathologist before making his final diagnostic decision on such tissues.

### 4.4 Evaluation of Accuracy Using ROC Curves

The ROC curves (Fig. 3) show the superior performance and accuracy of the present technique in differentiating between normal and diseased breast tissue samples. They also show that the present technique is highly accurate in differentiating between malignant and benign breast tissue samples, a result which supports the usefulness of the present technique in differentiating diseased breast tissue samples.

### 4.5 Molecular Basis of Origin of Differences Between Breast Tissues

One of the advantages of the FTIR technique is its ability to provide detailed quantitative information on the elements of the secondary structure of proteins (α-helix, ß-sheet, ß-turns, random coils, and aggregated strands) [29, 30, 46] through the deconvolution of the Amide I peak. Table 3 shows quantitatively that normal tissue maintains a high percentage of ordered secondary structure while diseased tissues are characterized by the presence of a greater percentage of unordered structures.

This study showed that it is possible to differentiate between three main breast tissue categories (normal, benign, and malignant) using the FTIR technique. The effect of the discrepancies underlying each category on the final FTIR profile and characterization parameters should be investigated in a future work.

## 5 Conclusion

FTIR is a reliable technique that can significantly differentiate between normal, malignant, and benign excised breast tissues. Among several characterization parameters that can be used to distinguish between the investigated tissues, the area under the peak at 2925 $cm^{-1}$ is the most sensitive (100 %) and most specific (100 %) parameter for differentiating normal from diseased breast tissue samples. The ratio of intensities (1657/2925 $cm^{-1}$) is the best parameter for differentiating between malignant and benign breast tissues (sensitivity = 100 % and specificity = 50 %). The ROC analysis shows that these parameters are also highly accurate. Moreover, the deconvolution of the Amide I peak reveals a higher percentage of ordered protein secondary structure elements (40.3 % α-helix and 57.3 % ß-sheet) in normal tissue compared to diseased tissues. The latter is characterized by an increase in unordered secondary structures (48 and 26.7 % unordered random coils in benign and malignant breast tissue samples, respectively). The present results strongly support the use of FTIR in conjunction with histopathological judgment for the characterization of excised breast tissue samples with high diagnostic performance.

#### Compliance with Ethical Standards

**Conflict of interest** The authors declare that they have no conflicts of interest.

## References

1. Mostaço-Guidolin, L. B., Murakami, L. S., Batistuti, M. R., Nomizo, A., & Bachmann, L. (2010). Molecular and chemical

characterization by Fourier transform infrared spectroscopy of human breast cancer cells with estrogen receptor expressed and not expressed. *Spectroscopy, 24*, 501–510.
2. Jemal, A., Bray, F., Center, M. M., Ferlay, J., Ward, E., & Forman, D. (2011). Global cancer statistics. *A Cancer Journal for Clinicians, 61*, 69–90.
3. Nystrom, L., Rutqvist, L. E., Wall, S., Lindgren, A., Lindqvist, M., & Ryden, S. (1993). Breast cancer screening with mammography: Overview of the Swedish randomised trials. *Lancet, 341*, 973–978.
4. Boyd, N. F., Guo, H., Martin, L. J., Sun, L., Stone, J., Fishell, E., et al. (2007). Mammographic density and the risk and detection of breast cancer. *The New England Journal of Medicine, 356*, 227–236.
5. Heywang-Köbrunner, S. H., Hacker, A., & Sedlacek, S. (2011). Advantages and disadvantages of mammography screening. *Breast Care, 6*, 199–207.
6. Lee, C. H., Dershaw, D. D., Kopans, D., Evans, P., Monsees, B., Monticciolo, D., et al. (2010). Breast cancer screening with imaging: Recommendations from the society of breast imaging and the ACR on the use of mammography, breast MRI, breast ultrasound, and other technologies for the detection of clinically occult breast cancer. *Journal of the American College of Radiology, 7*, 18–27.
7. Orel, S. G. (2000). MR imaging of the breast. *Radiologic Clinics of North America, 38*, 899–913.
8. Ci, Y. X., Gao, T. Y., Feng, J., & Guo, Z. Q. (1999). Fourier transform infrared spectroscopic characterization of human breast tissue: Implications for breast cancer diagnosis. *Applied Spectroscopy, 53*(3), 312–315.
9. Eckel, R., Huo, H., Guan, H. W., Hu, X., Che, X., & Huang, W. (2001). Characteristic infrared spectroscopic patterns in the protein bands of human breast cancer tissue. *Vibrational Spectroscopy, 27*, 165–173.
10. Krishna, C. M., Kurien, J., Mathew, S., Rao, L., Maheedhar, K., Kumar, K. K., et al. (2008). Raman spectroscopy of breast tissues. *Expert Review of Molecular Diagnostics, 8*(2), 149–166.
11. Meurens, M., Wallon, J., Tong, J., Noel, H., & Haot, J. (1996). Breast cancer detection by Fourier transform infrared spectrometry. *Vibrational Spectroscopy, 10*, 341–346.
12. Bigio, I. J., & Brown, S. G. (2004). Spectroscopic sensing of cancer and cancer therapy: Current status of translational research. *Cancer Biology & Therapy, 3*, 259–267.
13. Christoph, K., Larysa, S., Stephan, S. B., Gabriele, S., & Reiner, S. (2006). Identification of primary tumors of brain metastases by infrared spectroscopic imaging and linear discriminant analysis. *Technology in Cancer Research & Treatment, 5*(3), 291–298.
14. Kondepati, V. R., Heise, H. M., & Backhaus, J. (2008). Recent application of near Infra-red spectroscopy in cancer diagnosis and therapy. *Analytical and Bio-Analytical Chemistry, 390*(1), 125–139.
15. Movasaghi, Z., Rehman, S., & Rehman, I. U. (2008). Fourier transform infrared (FTIR) spectroscopy of biological tissues. *Applied Spectroscopy Reviews, 43*(2), 134–179.
16. Draux, F., Jeannesson, P., Gobinet, C., Sule-Suso, J., Pijanka, J., Sandt, C., et al. (2009). IR spectroscopy reveals effect of non-cytotoxic doses of anti-tumor drug on cancer cells. *Analytical and Bio-analytical Chemistry, 395*, 2293–2301.
17. Bellisola, G., & Sorio, C. (2012). Infrared spectroscopy and microscopy in cancer research and diagnosis. *American Journal of Cancer Research, 2*(1), 1–21.
18. Das, R. M., Ahmed, M. K., Mantsch, H. H., & Scott, J. E. (1995). FT-IR spectroscopy of methyl mercury-exposed mouse lung. *Molecular and Cellular Biochemistry, 145*, 75–79.
19. Gao, T., Feng, J., & Ci, Y. (1999). Human breast carcinomal tissues display distinctive FTIR spectra: Implication for the histological characterization of carcinomas. *Analytical Cellular Pathology, 18*, 87–93.
20. Gasper, R., Dewelle, J., Kiss, R., Mijatovic, T., & Goormaghtigh, E. (2009). IR spectroscopy as a new tool for evidencing antitumor drug signatures. *Biochimicaet Biophysica Acta, 1788*, 1263–1270.
21. Rehman, S., Movasaghi, Z., Darr, J. A., & Rehman, I. U. (2010). Fourier transform infrared spectroscopic analysis of breast cancer tissues; Identifying differences between normal breast, invasive ductal carcinoma, and ductal carcinoma in situ of the breast. *Applied Spectroscopy Reviews, 45*(5), 355–368.
22. Kong, J.,& Yu, S. (2007). Fourier transform infrared spectroscopic analysis of protein secondary structures. *Acta Biochimicaet Biophysica Sinica*, *39*(8), 549–559.
23. Janus, S. O., Malek, K. S., Glogowska, M. G., Walski, T., Komorowska, M., Witkiewicz, W., et al. (2012). Spectroscopic techniques in the study of human tissues and their components. Part I: IR spectroscopy. *Acta of Bioengineering and Biomechanics*, *14*(3), 101–115.
24. Yang, H., Yang, S., Kong, J., Dong, A., & Yu, S. (2015). Obtaining information about protein secondary structures in aqueous solution using Fourier transform IR spectroscopy. *Nature Protocols, 10*(3), 382–396.
25. Andrus, P. G. L., & Strickland, R. D. (1998). Cancer grading by Fourier transform infrared spectroscopy. *Biospectroscopy, 4*, 37–46.
26. Petibois, C., Gionnet, K., Gonçalves, M., Perromat, A., Moenner, M., & Déléris, G. (2006). Analytical performances of FT-IR spectrometry and imaging for concentration measurements within biological fluids, cells and tissues. *Analyst, 131*, 640–647.
27. Malins, D. C., Johnson, P. M., Barker, E. A., Polissar, N. L., Wheeler, T. M., & Anderson, K. M. (2003). Cancer-related changes in prostate DNA as men age and early identification of metastasis in primary prostate tumors. *Proceedings of the National Academy of Sciences, 100*(9), 5401–5406.
28. Prestrelski, S. J., Tedeschi, N., Arakawa, T., & Carpentert, J. F. (1993). Dehydration-induced conformational transitions in proteins and their inhibition by stabilizers. *Biophysical Journal, 65*, 661–671.
29. Haris, P. I., & Severcan, F. (1999). FTIR spectroscopic characterization of protein structure in aqueous and non-aqueous media. *Journal of Molecular Catalysis. B Enzymatic, 7*, 207–221.
30. Miller, L. M., Bourassa, M. W., & Smith, R. J. (2013). FTIR spectroscopic imaging of protein aggregation in living cells. *Biochimicaet Biophysica Acta, 1828*(10), 2339–2346.
31. Sasse, E. A. (2002). Objective evaluation of data in screening for disease. *Clinica Chimica Acta, 315*, 17–30.
32. Hada, T., Kondo, M., Yasukawa, K., Amuro, Y., & Higashino, K. (1999). Discrimination of liver cirrhosis from chronic hepatitis by measuring the ratio of Aleuria aurantia lectin-reactive serum cholinesterase to immune reactive protein. *Clinica Chimica Acta, 281*, 37–46.
33. Verma, A., Prasad, K. N., Singh, A. K., Nyati, K. K., Gupta, R. K., & Paliwal, V. K. (2010). Evaluation of the MTT lymphocyte proliferation assay for the diagnosis of neurocysticercosis. *Journal of Microbiological Methods, 81*, 175–178.
34. Elshemey, W. M., Mohamed, F. S., & Khater, I. M. (2013). X-ray scattering for the characterization of lyophilized breast tissue samples. *Radiation Physics and Chemistry, 90*, 67–72.
35. Grimes, D. A., & Schulz, K. F. (2002). Uses and abuses of screening tests. *Lancet, 359*, 881–884.
36. El Houseini, M. E., Mohammed, S. M., Elshemey, W. M., Hussein, T. D., Desouky, O. S., & Elsayed, A. A. (2005). Enhanced detection of hepatocellular carcinoma. *Cancer Control, 12*, 248–253.

37. Leeflang, M. M. G., Moons, K. G. M., Reitsma, J. B., & Zwinderman, A. H. (2008). Biasin sensitivity and specificity caused by data-driven selection of optimal cutoff values: Mechanisms, magnitude, and solutions. *Clinical Chemistry, 54*, 729–737.
38. Elshemey, W. M., Desouky, O. S., Fekry, M. M., Talaat, S. M., & Elsayed, A. A. (2010). The diagnostic capability of X-ray scattering parameters for the characterization of breast cancer. *Medical Physics, 37*, 4257–4265.
39. Swets, J. A. (1988). Measuring the accuracy of diagnostic systems. *Science, 24*, 1285–1293.
40. Greiner, M., Pfeiffer, D., & Smith, R. D. (2000). Principles and practical application of the receiver-operating characteristic analysis for diagnostic tests. *Preventive Veterinary Medicine, 45*, 23–41.
41. Landewe, R. B. M., & Heijde, D. M. F. M. (2003). Principles of assessment from a clinical perspective. *Best Practice and Research Clinical Rheumatology, 17*(3), 365–379.
42. Susi, H., & Byler, D. M. (1983). Protein structure by Fourier transform infrared spectroscopy: Second derivative spectra. *Biochemical and Biophysical Research Communications, 115*(1), 391–397.
43. Jørgensen, L., Vermehren, C., Bjerregaard, S., & Froekjaer, S. (2003). Secondary structure alterations in insulin and growth hormone water-in-oil emulsions. *International Journal of Pharmaceutics, 254*, 7–10.
44. Sarmento, B., Ferreira, D. C., Jorgensen, L., & Weert, V. M. (2007). Probing insulin's secondary structure after entrapment into alginate/chitosan nanoparticles. *European Journal of Pharmaceutics and Biopharmaceutics, 65*, 10–17.
45. Elshemey, W. M., Mohammad, I. A., & Elsayed, A. A. (2010). Wide-angle X-ray scattering as a probe for insulin denaturation. *International Journal of Biological Macromolecules, 46*, 471–477.
46. Barth, A. (2007). Infrared spectroscopy of proteins. *Biochimicaet Biophysica Acta, 1767*, 1073–1101.
47. Susi, H., Byler, D. M., & Purcell, J. M. (1985). Estimation of beta structure content of proteins by means of deconvolved FTIR spectra. *Journal of Biochemical and Biophysical Methods, 11*, 235–240.
48. Natalello, A., Ami, D., Brocca, S., Lotti, M., & Doglia, S. M. (2005). Secondary structure, conformational stability and glycosylation of a recombinant Candida rugosa lipase studied by Fourier-transform infrared spectroscopy. *Biochemical Journal, 385*, 511–517.
49. Iglesias, A., Arias, M., Santiago, P., Rodríguez, M., Mañas, J., & Saborido, C. (2007). Benign breast lesions that simulate malignancy: Magnetic resonance imaging with radiologic-pathologic correlation. *Current Problems in Diagnostic Radiology, 36*, 66–82.
50. Mersov, A., Mersov, G., Al-Ebraheem, A., Cornacchi, S., Gohla, G., Lovrics, P., et al. (2014). The differentiation of malignant and benign human breast tissue at surgical margins and biopsy using x-ray interaction data and Bayesian classification. *Radiation Physics and Chemistry, 95*, 210–213.
51. Yoshida, S., Miyazaki, M., Sakai, K., Takeshita, M., Yuasa, S., Sato, A., et al. (1997). Fourier transform infrared spectroscopic analysis of rat brain microsomal membranes modified by dietary fatty acids: Possible correlation with altered learning behavior. *Biospectroscopy, 3*(4), 281–290.
52. Wong, P. T. T., Papavassiliou, E. D., & Rigas, B. (1991). Phosphodiester stretching bands in the infrared spectra of human tissues and cultured cells. *Applied Spectroscopy, 45*, 1563–1567.
53. Wang, H. P., Wang, H. C., & Huang, Y. J. (1997). Microscopic FTIR studies of lung cancer cells in pleural fluid. *Science of the Total Environment, 204*, 283–287.
54. Mordechai, S., Mordechai, J., Ramesh, J., Levi, C., Huleihel, M., Erukhimovitch, V., et al. (2001). Application of FTIR microspectroscopy for the follow-up of childhood leukaemia chemotherapy. *Subsurface and Surface Sensing Technologies and Applications III, 4491*, 243–250.
55. Fukuyama, Y., Yoshida, S., Yanagisawa, S., & Shimizu, M. (1999). A study on the differences between oral squamous cell carcinomas and normal oral mucosas measured by Fourier transform infrared spectroscopy. *Biospectroscopy, 5*, 117–126.
56. Yang, Y., Sule-Suso, J., Sockalingum, G. D., Kegelaer, G., Manfait, M., & El Haj, A. J. (2005). Study of tumor cell invasion by Fourier transform infrared microspectroscopy. *Biopolymers, 78*, 311–317.
57. Chiriboga, L., Xie, P., Yee, H., Vigorita, V., Zarou, D., Zakim, D., et al. (1998). Infrared spectroscopy of human tissue. I. Differentiation and maturation of epithelial cells in the human cervix. *Biospectroscopy, 4*, 47–53.
58. Huleihel, M., Salman, A., Erukhimovich, V., Ramesh, J., Hammody, Z., & Mordechai, S. (2002). Novel optical method for study of viral carcinogenesis in vitro. *Journal of Biochemical and Biophysical Methods, 50*, 111–121.
59. Wu, J. G., Xu, Y. Z., Sun, C. W., Soloway, R. D., Xu, D. F., Wu, Q. G., et al. (2001). Distinguishing malignant from normal oral tissues using FTIR fiber-optic techniques. *Biopolymer (Biospectroscopy), 62*, 185–192.
60. Fung, M. F. K., Senterman, M. K., Mikhael, N. Z., Lacelle, S., & Wong, P. T. T. (1996). Pressure-tuning Fourier transform infrared spectroscopic study of carcinogenesis in human endometrium. *Biospectroscopy, 2*, 155–165.