Resonance Raman and Raman Spectroscopy for Breast Cancer Detection

www.tcrt.org
DOI: 10.7785/tcrt.2012.500325

Raman spectroscopy is a sensitive method to detect early changes of molecular composition and structure that occur in lesions during carcinogenesis. The Raman spectra of normal, benign and cancerous breast tissues were investigated in vitro using a near-infrared (NIR) Raman system of 785 nm excitation and confocal micro resonance Raman system of 532 nm excitation. A total number of 491 Raman spectra were acquired from normal, benign and cancerous breast tissues taken from 15 patients. When the 785 nm excitation was used, the dominant peaks in the spectra were characteristic of the vibrations of proteins and lipids. The differences between the normal and cancerous breast tissues were observed in both the peak positions and the intensity ratios of the characteristic Raman peaks in the spectral region of 700-1800 cm$^{-1}$. With 532 nm excitation, the resonance Raman (RR) spectra exhibited a robust pattern of peaks within the region of 500-4000 cm$^{-1}$. The intensities of four distinct peaks at 1156, 1521, 2854 and 3013 cm$^{-1}$ detected in the spectra collected from normal breast tissue were found to be stronger in comparison with those collected from cancerous breast tissue. The twelve dramatically enhanced characteristic peaks, including the enhanced amide II peak at 1548 cm$^{-1}$ in the spectra collected from cancerous breast tissue, distinguished the cancerous tissue from the normal tissue. Principal component analysis (PCA) combined with support vector machine (SVM) analysis of the Raman and RR spectral data yielded a high performance in the classification of cancerous and benign lesions from normal breast tissue.

Key words: Raman; Resonance Raman; Confocal micro Raman spectroscopy; Optical biopsy; Breast cancer; Amide I, II, III; Methyl; Methylene vibration bonds; Statistical methods: PCA; SVM; ROC.

Introduction

Breast cancer is one of the major causes of death in women. In 2012, approximately 226,870 new cases of invasive breast cancer and 39,510 new cases of non-invasive breast cancer in women were diagnosed in United States (US) (1). Breast cancer ranks second as a cause of cancer death in women after lung cancer (1).

*Corresponding authors:
R. R. Alfano, Ph.D.
Y. Zhou, M.D., Ph.D.
Y. Sun, Ph.D.
J. Y. Li, M.D., Ph.D.
L. X. Zhou, M.D.
S. Boydston-White, Ph.D.
V. Masilamani, Ph.D.
K. Zhu, B.S.
Yang Pu, Ph.D.
R. R. Alfano, Ph.D.*

1Physics Department, Institute for Ultrafast Spectroscopy and Lasers, The City College of New York, New York, NY 10031, USA
2The General Hospital of the Air Force, No.30 Fuchenglu, Haidian District, Beijing 100142, China
3Electrical Engineering Department, The City College of New York, New York, NY 10031, USA
4Beijing Cancer Hospital, No.52 Fuchenglu, Haidian District, Beijing 100142, China
5Science Department, City University of New York, Borough of Manhattan Community College, New York, NY 10007-1097, USA

Abbreviations: Resonance Raman (RR); Near-Infrared (NIR); Invasive Ductal Carcinoma (IDC); Ductal Carcinoma In Situ (DCIS); Principal Component Analysis (PCA); Support Vector Machine (SVM); Receiver Operating Characteristic (ROC).
Early diagnosis is the key to increase the five-year survival rate of cancer patients.

The application of needle aspiration biopsy for the diagnosis of neoplastic conditions attracted attention in the early 20th century. Two methods of nonsurgical investigation of breast tissue were studied in the 1920’s, namely transillumination (with flash light) in 1929 and radiography (2, 3). The first concerted effort to use needle aspiration and biopsy for the diagnosis of cancer was undertaken at Memorial Hospital in New York in 1929 by Cutler (2). The modern stereotaxic core biopsy instruments were introduced in the 1980’s. Breast cancer screening and diagnosis are now performed by a combination of the physician’s manual search for palpable lesions – X-ray mammography, ultrasonography (US), X-ray computed tomography (CT), optical coherence tomography (OCT), magnetic resonance imaging (MRI) and positron emission tomography (PET) then followed by one or more tissue biopsies. 70-90% of the suspicious lesions detected by mammography are determined to be benign lesions upon biopsy. However, 20% of malignant lesions go undetected by mammography (4-6). Minimally invasive or non-invasive optical spectroscopy techniques for the diagnosis of suspicious lesions in real-time could reduce patient trauma and discomfort, time to diagnosis and the high medical costs associated with traditional biopsy (7).

As early as 1984, a renaissance occurred. Alfano and coworkers and students pioneered the field of optical biopsy using native fluorescence (autofluorescence) spectroscopy for probing electrons states (8, 9). The first briefly reported RR spectra of human breast tissue were measured by Alfano in 1987 using 457 and 480 nm excitations (9). Later in 1991, a detailed investigation of Raman spectroscopy was used to probe the vibrational states by Alfano and coworkers in spectra collected from normal and cancerous tissues using near-infrared (NIR) 1064 nm excitation (10). Raman spectroscopy provides detailed information about the biochemical composition of tissues from the vibrational states found in the spectra and information for diagnosing disease from the changes associated with those vibrational states. The other techniques of optical biopsy utilize spectroscopy that includes intrinsic native fluorescence, Stokes Shift and Raman spectroscopy to identify neoplastic transformations in cells and tissues on the molecular level, in various organs sites in near real-time, \textit{ex vivo} and \textit{in vivo} (11-14). The Raman spectrum is more robust than the other methods because of the sharp peaks associated with different vibrational modes.

The Raman effect is named after C. V. Raman who discovered the effect in 1928. Over the past three decades, there has been tremendous growth worldwide of optical spectroscopic techniques (15-24) using 785 nm laser light from diode lasers, especially in Raman spectroscopic characterization and discrimination of breast malignancy (6, 18, 19, 25-27). However, few works disclose resonance Raman spectra differences of the cancerous and normal tissues and study biological basis of these differences (12). The resonance effect in RR spectroscopy occurs when the energy of the incoming photon approaches the energy level of an electronic transition of the electrons associated with the bonds in the molecules. When the frequency of the laser beam overlaps an electronic transition energy level, the vibrational modes of that molecule’s transition are greatly increased and the resulting spectra exhibit Raman scattering intensity enhancement. For example, the light resonance effect at an excitation of 532 nm enhances the vibrational modes associated with the bonds in flavin molecules, while other vibrational modes remain unaffected. The resonance enhancement of the vibrational modes in RR spectroscopy becomes useful for studying large biomolecules with chromophores embedded in their structure and the vibrational modes associated with the absorbing groups. This allows RR spectra to be collected from small molecular concentrations and particular spectral patterns may emerge and may be utilized as markers in the vibrational spectra. The resonance effect is the major difference between the spectra collected using conventional Raman spectroscopy and RR spectroscopy – with the spectra collected using RR Raman exhibiting smaller number of peaks, but peaks with higher resolution. Previous non-allowed peaks can be recognized by targeting specific electronic transitions in key biomolecules in tissues, cells and organelles such as flavins, NADH, collagen, elastin, mitochondrion, cytochromes and other heme proteins.

In comparison to the conventional Raman spectroscopy, RR spectroscopy can:

1. Better detect molecular vibrations in biological tissues, affected by the structure, composition, environment and molecular interactions in cells and tissues in their natural state without adding labels or contrast enhancing agents.
2. Detect some infrared-active (IR-active) vibrations whose signals are too weak to be detected in the conventional Raman spectra (16, 28).
3. Detect more modes enhanced of vibrational modes, due to the excitation wavelength close to or overlapping with the absorption wavelength of molecules, that are key to the characterization of molecules within cells and tissues.
4. Detect biological changes on the molecular level with a higher resolution.
5. Transcend the limitations of the standard absorption–emission processes, limitations due to the large density of vibrational peaks, broadened by the contributions from homogeneous and inhomogeneous biomedical processes, to produce a higher resolution of spectra.
6. Target vibrational groups associating with biomolecules (tryptophan, collagen, elastin and NADH) in tissues and cells.

This study investigates more extensively the differences in the resonance Raman (RR) and regular Raman (RS) spectra collected from normal and cancerous breast tissues using 532 and 785 nm excitation wavelengths. The thrust of this report is to demonstrate the importance of RR in cancer research. This work extends previous works in breast cancer detection from conventional Raman spectroscopy into resonance Raman.

Materials and Methods

Tissue Acquisition

Human normal and malignant breast tissue samples, exhibiting varying stages of disease were obtained from the National Disease Research Interchange (NDRI), the Cooperation Human Tissue Network (CHTN) and the Beijing Cancer Hospital (Beijing, China). The tissue specimens were as follows:

- All breast tissues specimens including normal, benign and cancerous tissues were from female patients.
- All cancer tissue specimens were invasive ductal carcinoma (IDC), but two of the cancer tissues were ductal carcinoma in situ (DCIS).
- The benign breast lesion tissues were designated as fibroadenoma with necrosis (in the pathology report, meaning that necrotic breast cancer cells can be seen within the tissue sample).
- The bulk of the tissue came from patients ranging in age from 32 to 71 years, with the median age of 55 years, with one normal tissue specimen from a female of 16 years.

Although the benign breast lesions represent an abnormal proliferation, the morphology and composition of this tissue are more similar to that of normal breast tissue and these lesions will not become cancerous. In comparison, cancerous breast tissue exhibits gross morphological differences, differ in composition and yielded spectral features that are markedly different.

This study were exempt from Institutional review board (IRB) review, as the tissues that were not obtained from tissue banks, were collected and maintained in such a manner that subjects cannot be identified, directly or through identifiers linked to the subjects, which falls under Exemption #4. Only anonymous tissue specimens acquired from a tissue bank (NDRI) without patient identifiers was used in this study. Whenever possible, matched normal controls were obtained from the same patient, if available, from the tissue bank.

Tissue Preparation

Specimens were not chemically treated prior to the spectroscopic studies. The tissue specimens were frozen using liquid nitrogen (LN2, Snap Frozen) and maintained frozen until they were packed in dry ice and shipped. All tissues arrived still frozen on dry ice, uncut and in irregular shapes and were maintained at −80°C until they were thawed to ambient room temperature for the spectroscopic studies. In preparation for the collection of RR spectra with an excitation wavelength of 785 nm, specimens of random shapes were mounted inside standard quartz cuvettes (1 × 1 × 4.5 cm) NSG Company (USA). Care was taken to minimize gaps between the tissue and the inside walls of the cuvette. In preparation for the collection of spectra using the confocal micro Raman system with 532 nm excitation, the specimens were placed on a quartz plane. The average size of the specimens was approximately 10 × 20 × 2 mm.

Raman Spectral Measurements

A total of 436 Raman spectra were acquired from specimens of normal, benign and cancerous tissues taken from 10 patients using the R2000 Raman system (Ocean Optics Inc., Dunedin, FL, USA) and 55 RR spectra were acquired from normal and cancerous tissue specimens taken from three patients – each set taken from the same patient – using the confocal micro RR system.

NIR Regular Raman Spectra Acquisition (785 nm)

The Raman spectra from normal, benign and the carcinoma tissues were acquired using Ocean Optics R-2000 NIR-Raman spectrometer. This Raman system consists of multimode solid-state diode laser with an output power of 500 mW at room temperature and an “all-in-one” fiber optic probe with a spectral resolution of 15 cm⁻¹. At the focal point in the tissue specimen, the laser power was 175 mW, with an excitation spot size 0.5 mm. The Raman spectra were acquired in the backscattered direction (29), with an integration time of 30 seconds. Three spectra were collected at each location and averaged to reduce the noise level.

Confocal Micro RR Spectra Acquisition (532 nm)

The RR spectra were acquired using the Jobin Yvon (JY HR-800, France), confocal micro Raman spectrometer, equipped with a multichannel modular triple system and confocal microscopy with a 100X objective lens. The excitation laser spot diameter focused on the sample position was about 1 μm. The 532 nm wavelength solid-state Coherent Company-Verdi-2 diode laser (Coherent Company, Santa Clara, CA, USA) with an output of 3.5 mW was used as an excitation...
source. The laser power at the sample was kept to 0.9 mW and the typical spectrum acquisition time was 20-40 seconds, with an integration time of 4 × (20-40) seconds. All spectra were acquired at room temperature. The final spectral resolution was 2 cm⁻¹ in the range of interest, from 500 to 4000 cm⁻¹.

Data Analysis: Spectrum Pre-processing

The spectra collected using Ocean Optics R-2000 Raman spectrometer yielded a strong fluorescence background. This background was removed by using the Microcal Origin Pro 7.0 software (Microcal, Inc., Northampton, MA) and the fluorescence background was removed by fitting the spectra to a quadratic function. Although the quadratic fluorescence background correction is not justified theoretically, it greatly simplifies computation and ensures high performance in the cancer detection in the experiment. The spectra collected using the RR system demonstrated almost no fluorescence background due to the two active confocal pinholes that eliminate fluorescence light and stray light, but the software was used to remove the background, if needed.

Analysis Method

The cancerous, benign and normal tissue specimens were classified by the support vector machine (SVM) which is well-known as one of the most powerful classifiers. Since the Raman and RR spectra are practically sampled into high dimension in the Raman shifts, training of SVM using directly the high dimensional data might be suffered from the curse of dimensionality. To reduce significantly the data dimensionality while retaining the most power of the data, the data spectra can be projected onto a subspace spanned by a few of principal components (PCs) along which the data present the most power, so called principal component analysis (PCA) (30-32). Specifically, let $X = (x(1), \ldots, x(N))$ be the matrix of Raman spectra of all $N$ samples, which can be decomposed into

$$X = USV^T,$$  \[1\]

where $S = \text{diag}(s_1, \ldots, s_N)$ is the diagonal matrix of singular values in the descending order, $U = (u(1), \ldots, u(N))$ and $V = (v(1), \ldots, v(N))$ are the matrices of column and row singular vectors, and $s_i^2$ is the power of data in the dimension of $i$th singular vector. In general, a small number $n$ of column singular vectors $u(1), \ldots, u(n)$, the PCs, can be chosen such that most power of data are located in the subspace spanned by them. The significance of the selected PCs can be evaluated by the power usage, defined as the portion of power located in the subspace spanned by the selected PCs to the total power of data, which can be calculated by

$$P = \frac{\sum_{i=1}^n s_i^2}{\sum_{i=1}^N s_i^2}.$$  \[2\]

In practice, $n$ can be much smaller than the dimensionality of data. All Raman spectra are projected onto the subspace and the dimensionality is reduced to $n$. Since most power of data is retained in this subspace, the reduction of dimensionality only slightly reduces the discriminability of data in cancer detection. On the other hand, the significant reduction of dimensionality can effectively obviate the curse of dimensionality in training of the SVM classifier

$$w^Ty = b,$$  \[3\]
a hyperplane in the subspace. The sensitivity and specificity can be obtained by the SVM classifier. By adjusting the distance $b$ of the classifier to the origin and keeping the normal direction determined by $w$, the receiver operating characteristic (ROC) in terms of sensitivity and specificity can be obtained. The ROC curve is a graphical plot of true positive rate vs. false positive rate. Accuracy of the performance can be measured by the area under the ROC curve (AUC).

Results and Discussion

NIR Regular Raman Spectroscopy

The Raman spectra are characterized by six prominent peaks at 821, 1080, 1262, 1318, 1453 and 1663 cm⁻¹. Figure 1A, 1B and 1C display three Raman spectral curves obtained in vitro, from normal breast tissues, benign lesions and cancerous breast tissues, respectively, using an excitation wavelength of 785 nm. Each spectrum was labeled according to the pathological diagnosis and spans the scan region of 700-1800 cm⁻¹.

The three major Raman peaks investigated in this study were at 1663 cm⁻¹, the protein amide I peak due to carbonyl (C=O) stretching; the 1453 cm⁻¹, the protein amide II peak due to methylene (−CH₂−) deformation; and the 1262 cm⁻¹, protein amide III peak due to C−N single bond stretching and N−H bending. A lipid peak at 1746 cm⁻¹ in the spectra from the cancerous tissues (Figure 1C), was found to have been greatly diminished compared with Raman spectra from normal tissues in Figure 1A and the benign lesions in Figure 1B. The differences in the spectra of the normal (Figure 1A) and cancerous tissues (Figure 1C) in the intensity ratio of the peaks at 1453 to 1663 cm⁻¹ is striking.

The spectrum in Figure 1C displays a higher intensity ratio (1663-1453 cm⁻¹) compared with the spectrum in Figure 1A. In the spectra from the cancerous tissue, the intensity of the amide I peak around 1663 cm⁻¹ appears stronger than 1453 cm⁻¹ peak, which implies an increase of protein concentration.

The typical Raman spectrum from benign lesions (Figure 1B), shows spectral peaks more similar to the spectrum from normal tissues (Figure 1A) than to that of spectrum collected...
Resonance Raman Spectroscopy Detection of Breast Cancer

from cancerous tissue (Figure 1C), especially in the peak of 1445 cm\(^{-1}\) (fatty acids–lipid protein–CH\(_2\)) and the 1662 cm\(^{-1}\) (amide I) peaks and the ratios of those peak intensities. When one considers the protein peak at 1607 cm\(^{-1}\) is stronger than the fatty acid methylene (–CH\(_2\)) bond vibration at 1445 cm\(^{-1}\) in the spectra collected from benign lesions (Figure 1B), this pattern more closely resembles the pattern of the spectra collected from cancerous tissue, that showed an increase in the protein contribution due to the differences in the connective extracellular matrix (stroma) in tumor tissues. This is indicative of the capabilities of using Raman spectra in distinct classification of the three types of clinically encountered tissues. The classification method and results are presented.

The results showed a clear distinction between the Raman spectra collected from normal and cancerous tissues in the spectroscopic region of 700-1800 cm\(^{-1}\), by the peak positions and the peak intensity ratios of the spectral peaks characteristic of the cellular components (5, 6, 10, 15).

Confocal Micro RR Spectroscopy

The RR spectral experiments, carried out using JY HR-800, Confocal Raman micro-spectrometer with excitation wavelength of 532 nm, have been demonstrated as useful in numerous applications for the detection of proteins, nucleic acids and polycyclic aromatic hydrocarbons. Using the 532 nm excitation, generally results in a larger number of Raman spectral peaks enhanced because the 532 nm excitation approaches the absorption wavelengths of more species of cellular molecules, such as some proteins, enzymes, coenzymes and lipids. In addition, pre-resonance Raman enhancements also dramatically increased the number of peaks in the Raman spectra (33-35). In this study, the RR spectra collected showed a limited number of sharp and very intense peaks resulting from resonance enhancements.

Figure 2 displays the RR spectra collected from tissues diagnosed as normal breast tissue (normal) and stage II IDC cancer, in the scan region of 500-4000 cm\(^{-1}\). Spectral peaks near...
1545, 1585 and 1605 cm\(^{-1}\) as displayed in Figures 2 and 3, with the amide II peak (1545 cm\(^{-1}\)) were greatly enhanced in the spectra collected from the cancerous tissues.

Figure 3 displays the RR spectra from normal breast tissue and cancerous breast tissue in an enlarged scale region of 500-1800 cm\(^{-1}\). The spectra collected from normal breast tissue exhibited seven major peaks – 750, 1004, 1156, 1306, 1441, 1521 and 1656 cm\(^{-1}\). Figure 4 exhibits intense peaks in the 2854-3013 cm\(^{-1}\) region. In making assignments, the peak at 750 cm\(^{-1}\) corresponds to the C–H out-of-plane deformations, the methyl rocking modes and C–C stretching modes are between 1000 and 1200 cm\(^{-1}\); the in-plane C–H bends, methyl (–\(\text{CH}_3\)) deformations and C=C stretching modes are between 1200 and 1600 cm\(^{-1}\); and the region between 2800 and 3100 cm\(^{-1}\) is quite complex, containing at least ten Raman vibration peaks, assigned to methyl (–\(\text{CH}_3\)) and methylene (–\(\text{CH}_2\)) groups and the Fermi resonance vibrations of the \(\delta\)(C–H) overtones (28, 36-38). Specific assignments of individual peaks can be found in Table I.

The RR spectra collected twelve enhanced characteristic peaks from the cancerous breast tissue, including an enhanced amide II peak at 1548 cm\(^{-1}\) within the region of 500-1800 cm\(^{-1}\) (Figures 2 and 3) and a broad band of peaks at 2807-2986 cm\(^{-1}\) within the 2500-3500 cm\(^{-1}\) region was observed (Figures 2 and 4), due to the methyl (–\(\text{CH}_3\)) and methylene (–\(\text{CH}_2\)) group bonds, are markedly diminished in the spectra of cancerous tissues vs. the spectra collected from the normal tissues (39-43).

In comparing the RR spectra of the cancerous tissues vs. normal tissues, displayed in Figure 3, it is important to note that the amide II peak at 1548 cm\(^{-1}\) – a mixture of the C–N stretching motion and the C–N–H angle bending motion – is enhanced in the cancerous tissue, a peak in the 1087 cm\(^{-1}\) area (sugar, C–OH), 1172 cm\(^{-1}\) (methyl, –\(\text{CH}_3\) rock), 1223 cm\(^{-1}\) (enhanced amide III) and in the 1330-1380 cm\(^{-1}\) range, the peaks are enhanced, representing the vibrations due to the methyl (–\(\text{CH}_3\)) and methylene (–\(\text{CH}_2\)) groups (36-38). Although the methyl (–\(\text{CH}_3\)) 1378 cm\(^{-1}\) deformation peak is enhanced in the spectra of cancerous tissue, Figure 4 shows that the overtone at 2727 cm\(^{-1}\) is diminished when compared to the spectra of normal tissue and a relative decrease in lipid concentration, exhibited by the diminished peaks in the 2950-3050 cm\(^{-1}\) range in the spectra collected from cancerous tissues (Figures 2 and 4).

The peak near 2935 cm\(^{-1}\) (Figure 4), due to the antisymmetric stretch of methylene (–\(\text{CH}_2\)) groups and a possible contribution of an overtone of the in-phase symmetric methyl (–\(\text{CH}_3\)) stretching vibration and Fermi resonance with the C–H vibration at 1460 cm\(^{-1}\) deformation peak overtone, though this vibration yields a diminished peak in the spectra of cancerous tissue.

These changes can be used as a mode marker for discriminating the spectra from cancerous tissues from that from normal tissues such as the 1358 cm\(^{-1}\) peak, from the \(\text{CH}_3\)\(-\text{(C=O)}\) stretch; 1378 cm\(^{-1}\), from the methyl (–\(\text{CH}_3\)) in-phase deformation; the 1463 cm\(^{-1}\) peak, from the antisymmetric methyl (–\(\text{CH}_2\)) out-of-plane deformation; the 1548 cm\(^{-1}\) peak, from the protein amide II derived mainly from the in-plane N–H bending; the 1589 cm\(^{-1}\) peak, the C–C stretching and phosphate (O–P–O) symmetric stretching vibrational mode, characteristic of nucleic acids; the 1605 cm\(^{-1}\) peak, C–O stretching, and C=C bending in phenylalanine and tyrosine, in proteins; the 1639 and 1656 cm\(^{-1}\) peaks, the amide I peak of \(\alpha\)-helical and unordered proteins (44-47); and the 1746 cm\(^{-1}\) peak, diminished in the spectra from cancerous cells, from the carbonyl (C–O) stretch of phospholipids (18, 34, 35). The amide II peak at about 1550 cm\(^{-1}\) is due to
Resonance Raman Spectroscopy Detection of Breast Cancer

out-phase N–H bending over C–N and is usually weak but can be observed in the RR (48, 49).

In the spectra collected from normal tissue, an observed methylene (−CH₂) peak grows sharply at 2854 cm⁻¹ as seen in Figure 4, while in the spectra from cancerous tissues the peaks in that same area appear to be broader, less intense and less distinct. This difference in bandwidth and diminished intensity in the spectra of cancerous tissues vs. the spectra of normal tissues may be the result of the fact that the concentration of adipose cells is lower in cancerous breast tissues, as opposed to normal breast tissues, although changes in the patterns of DNA and protein methylation have been implicated in tumorigenesis. Although changes in methylation been detected in cancerous cells and tissues in some other cancers, the dramatically lower concentration of adipose cells in breast cancer tissues, thus a greatly diminished spectral contribution due to those vibrations, could mask any detection by vibrational spectroscopy of changes in the methylation states of DNA and/or protein that could be disease-related (43, 50).

<table>
<thead>
<tr>
<th>Raman shift cm⁻¹</th>
<th>Assignment</th>
<th>Intensity changes in lesion compare to normal tissue</th>
</tr>
</thead>
<tbody>
<tr>
<td>676</td>
<td>νδ (CCN), Vinyl &amp; Porphyrin modes</td>
<td>Enhanced</td>
</tr>
<tr>
<td>750</td>
<td>CH₂ rock, symmetric breathing of tryptophan</td>
<td>Enhanced</td>
</tr>
<tr>
<td>975</td>
<td>=C−H out-of-plane deformation</td>
<td>Enhanced</td>
</tr>
<tr>
<td>1004</td>
<td>Symmetric CC aromatic ring breathing</td>
<td>Similarity</td>
</tr>
<tr>
<td>1098-1130</td>
<td>CC stretch, CC skeletal stretch trans, PO₂ symmetric</td>
<td>Enhanced</td>
</tr>
<tr>
<td>1156</td>
<td>CH₁ rock</td>
<td>Distinct peak in normal</td>
</tr>
<tr>
<td>1172</td>
<td>CH₁ rock</td>
<td>Similarity</td>
</tr>
<tr>
<td>1223</td>
<td>Amide III</td>
<td>Enhanced</td>
</tr>
<tr>
<td>1306</td>
<td>Amide III (N−H), α-helix, C−C Str &amp; C−H bending</td>
<td>Enhanced</td>
</tr>
<tr>
<td>1338</td>
<td>CH₁ deformation</td>
<td>Enhanced</td>
</tr>
<tr>
<td>1358</td>
<td>CH₁−(C=O), strong than hydrocarbons</td>
<td>Enhanced</td>
</tr>
<tr>
<td>1421-1463</td>
<td>δ(CH₂) bending, δ(CH₁) out-of-phase deformation</td>
<td>Similarity or weaker</td>
</tr>
<tr>
<td>1521</td>
<td>Amide II, shift to 1548</td>
<td>Distinct peak in normal</td>
</tr>
<tr>
<td>1548</td>
<td>Amide II, in plane N−H bending</td>
<td>Enhanced</td>
</tr>
<tr>
<td>1589</td>
<td>C−C stretching, C−H bending</td>
<td>Enhanced</td>
</tr>
<tr>
<td>1605</td>
<td>CO stretching, C=C bending</td>
<td>Enhanced</td>
</tr>
<tr>
<td>1639</td>
<td>Amide I in α-helix</td>
<td>Shit to 1639 cm⁻¹</td>
</tr>
<tr>
<td>1656</td>
<td>Amide I unordered</td>
<td>Reduced</td>
</tr>
<tr>
<td>1746</td>
<td>ν(C=O)</td>
<td>Reduced</td>
</tr>
<tr>
<td>2277</td>
<td>1378 cm⁻¹ bend overtone</td>
<td>Reduced</td>
</tr>
<tr>
<td>2807</td>
<td>Aliphatic −N−CH₃, amine</td>
<td>Weaker</td>
</tr>
<tr>
<td>2854</td>
<td>ν(CH₃)</td>
<td>Distinct peak in normal</td>
</tr>
<tr>
<td>2892</td>
<td>ν(CH₃, FR)</td>
<td>Distinct peak</td>
</tr>
<tr>
<td>2900</td>
<td>CH₃ strong symmetric str.</td>
<td>Weaker</td>
</tr>
<tr>
<td>2933-2935</td>
<td>ν(CH₃, FR)</td>
<td>Distinct peak</td>
</tr>
<tr>
<td>2985</td>
<td>CH₃ medium asymmetric str.</td>
<td>Weaker, distinct peak</td>
</tr>
<tr>
<td>3103</td>
<td>CH₃ (C=O)</td>
<td>Distinct peak in normal</td>
</tr>
<tr>
<td>3166</td>
<td>ν(O−H) water band</td>
<td>Increased</td>
</tr>
<tr>
<td>3227</td>
<td>ν(O−H) water band</td>
<td>Increased</td>
</tr>
</tbody>
</table>

ν: Stretch; FR: Fermi resonance; δ: Bending; Refs. [29, 36-38, 47, 48].

Table I

Raman band frequencies and assignments of breast tissue.

Statistical Analysis

The NIR-Raman spectra of 103 normal tissues, 38 benign lesions and 150 cancerous tissues were analyzed using the SVM and PCA. The SVM with PCA in the analysis of Raman spectra is highly effective for cancer detection. As shown in Table II, the performance is improved as the power of the data located in the subspace spanned by the selected PCs increases. When 16 of the most important PCs were used, the total power ratio of the data in the subspace over the entire data set reaches 99.6% and the AUC achieves 99.4% or higher. Moreover, the sensitivity and specificity are 0.968 and 1.000 respectively, when compared to the pathology reports – the “gold standard” – for the characterization of the spectra from cancerous tissues against spectra from benign lesions (as a class) against the spectra from normal tissues (as a class), while the sensitivity and specificity are 0.980 and 0.979 respectively, for the characterization of the spectra from cancerous tissues (as a class) against the spectra from benign lesions and normal tissues (as a class). Increasing the number of PCs over the 16 PCs no longer improved the performance.
tissues, projected onto the subspace spanned by the PC 1 and PC 2 as diagnostically significant components. The black dotted line is the SVM classifier discriminating the spectra from the benign breast lesions and cancerous breast tissues (as a class) from the spectra from normal breast tissues (as a class) that yielded a diagnostic sensitivity of 0.872 and specificity of 0.612 (given in Table II). The cyan dashed line is the SVM classifier discriminating the spectra from the cancerous breast tissues (as a class) from the spectra from the benign lesions and normal breast tissues (as a class), that yielded a diagnostic sensitivity of 0.633 and a specificity of 0.681 (given in Table II).

Figure 7 displays the NIR-Raman spectra from the normal breast tissues, benign breast lesions and DCIS breast cancer tissues, projected onto the subspace spanned by the PC 1 and PC 4 as diagnostically significant components. The black dotted line is the SVM classifier discriminating the spectra from the benign breast lesions and cancerous breast tissues (as a class) from the spectra from normal breast tissues, that yielded a diagnostic sensitivity of 0.894 and 0.767 (not given in Table II). The cyan dashed line is the SVM classifier discriminating the spectra from the cancerous breast tissues (as a class) from the spectra from benign lesions and normal breast tissues (as a class), that yielded a diagnostic sensitivity of 0.813 and 0.766 (given in Table II).

Figure 8 displays the ROC under the PCA analysis and SVM classifier vs. the number of PCs as the diagnostically significant component(s) in discriminating the cancerous and benign breast tissues (as a class) from the normal breast tissues.

Figure 9 displays the ROC under the PCA analysis and SVM classifier vs. the number of PCs as the diagnostically significant component(s) in discriminating the cancerous breast tissues from the benign lesions and the normal breast tissues.

Table II
Classification performance in analysis of Raman spectra.

<table>
<thead>
<tr>
<th>P(%)</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>AUC</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal vs. benign and cancerous</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 PC</td>
<td>0.659</td>
<td>0.856</td>
<td>0.505</td>
<td>0.664</td>
<td>0.800</td>
<td>0.576</td>
</tr>
<tr>
<td>2 PCs</td>
<td>0.897</td>
<td>0.872</td>
<td>0.612</td>
<td>0.773</td>
<td>0.633</td>
<td>0.773</td>
</tr>
<tr>
<td>3 PCs</td>
<td>0.948</td>
<td>0.878</td>
<td>0.670</td>
<td>0.845</td>
<td>0.860</td>
<td>0.864</td>
</tr>
<tr>
<td>4 PCs</td>
<td>0.964</td>
<td>0.941</td>
<td>0.825</td>
<td>0.950</td>
<td>0.840</td>
<td>0.930</td>
</tr>
<tr>
<td>8 PCs</td>
<td>0.989</td>
<td>0.979</td>
<td>0.932</td>
<td>0.990</td>
<td>0.933</td>
<td>0.970</td>
</tr>
<tr>
<td>16 PCs</td>
<td>0.996</td>
<td>0.968</td>
<td>1.000</td>
<td>0.994</td>
<td>0.980</td>
<td>0.997</td>
</tr>
<tr>
<td>Normal and benign vs. cancerous</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 PC</td>
<td>0.664</td>
<td>0.800</td>
<td>0.576</td>
<td>0.633</td>
<td>0.840</td>
<td>0.930</td>
</tr>
<tr>
<td>2 PCs</td>
<td>0.773</td>
<td>0.633</td>
<td>0.773</td>
<td>0.860</td>
<td>0.840</td>
<td>0.930</td>
</tr>
<tr>
<td>3 PCs</td>
<td>0.845</td>
<td>0.860</td>
<td>0.864</td>
<td>0.840</td>
<td>0.840</td>
<td>0.930</td>
</tr>
<tr>
<td>4 PCs</td>
<td>0.950</td>
<td>0.840</td>
<td>0.930</td>
<td>0.840</td>
<td>0.840</td>
<td>0.930</td>
</tr>
<tr>
<td>8 PCs</td>
<td>0.990</td>
<td>0.933</td>
<td>0.970</td>
<td>0.933</td>
<td>0.933</td>
<td>0.970</td>
</tr>
<tr>
<td>16 PCs</td>
<td>0.994</td>
<td>0.980</td>
<td>0.997</td>
<td>0.980</td>
<td>0.980</td>
<td>0.997</td>
</tr>
</tbody>
</table>
The RR spectra collected from a small set of 10 normal and 5 cancerous tissues were also analyzed by the PCA and SVM. The results are presented in Table III and Figures 10-12. When three PCs were used, the cancerous tissues were already detected with both sensitivity and specificity at 100%. To obtain a statistically more convincing result in RR data analysis, a larger set of tissues will be considered in future work.
Figure 10 displays the first four principal components obtained by PCA for the RR spectra collected from normal breast tissues and IDC breast cancer tissues.

Figure 11 displays the RR spectra from normal breast tissues and IDC breast cancer tissues projected onto the subspace spanned by the PC 1 and PC 2 as diagnostically significant components. The black dotted line is the SVM classifier discriminating cancerous tissues from the normal tissues that yielded a diagnostic sensitivity of 0.80 and specificity of 0.90, as compared with the pathology results, given in Table III.

Figure 12 displays the RR spectra from normal breast tissues and IDC breast cancer tissues projected onto the subspace spanned by the PC 1 and PC 3 as diagnostically significant components. The black dotted line is the SVM classifier discriminating cancerous tissues from the normal tissues that yielded a diagnostic sensitivity of 1.00 and specificity of 1.00 (not given in Table III).

Conclusion

This study shows the capabilities of the use of Resonance Raman spectroscopy to classify three types of clinically encountered tissues – normal breast tissue, benign lesions and cancerous breast tissue. The spectra collected from tissues using RR spectroscopy revealed new information. The methyl group peaks in the spectra collected from breast cancer tissue have been identified with enhanced intensity at 1378 cm\(^{-1}\), and another methyl group on an aromatic ring that has one peak near 2935 cm\(^{-1}\) is due to the in-phase symmetric methyl (–CH\(_3\)) stretching vibration and a vibration in Fermi resonance with the –CH\(_3\) 1460 cm\(^{-1}\) deformation peak overtone – a weaker vibration in the spectra collected from breast cancer tissues. Multivariate statistical analyses were carried out of the Raman spectral data to develop a classification model to discriminate the spectra from abnormal breast tissues from the spectra from normal breast tissues.

The RR spectral analyses of the methyl group vibrations in the peaks at 1378 cm\(^{-1}\), the broadened band of peaks at 2854-2985 cm\(^{-1}\) and the intense amide II peak, as well as another eleven enhanced peaks may be useful in distinguishing the spectra collected from breast cancer tissues from the spectra collected from normal breast tissues. RR spectroscopy detects biochemical changes within cells and tissues during biological processes, including the progression of disease states.

The intense resonance enhancement displayed in two groups of peaks – one at 1338-1378 cm\(^{-1}\) and the second group at 1548-1605 cm\(^{-1}\) – suggests that the 532 nm excitation wavelength matched (or closely matched) the molecular absorption wavelengths for molecular compounds in the cells and tissues. For example, the metalloprotein, hemoglobin, has one absorption band at 534 nm. Similarly, the mitochondrial electron transport protein, cytochrome c, has one absorption wavelength at 552 nm (under hypoxia conditions) and a two-photon absorption (TPA) process under RR conditions. The RR spectroscopic technique, using the 532 nm wavelength excitation, revealed key differences in the spectra from cancerous breast tissues vs. normal breast tissues, in peaks associated with proteins on the molecular level in cancer.
Resonance Raman Spectroscopy Detection of Breast Cancer

concerning the manuscript. Mr. C. Y. Zhang, Mr. M. Z. Fan, and Dr. X. H. Ni for their ordered and kept four normal and cancerous breast tissue authors acknowledge the help of Ms. L. A. Sordillo for helping of #W81XWH-11-1-0335 (CUNY RF # 47204-00-01). The This research is supported in part by US Army Medical reporting or to the writing of the paper.

Conflict of Interest

We certify that regarding this paper, no actual or potential conflicts of interests exist. The paper is original, has not been accepted for publication nor is concurrently under consideration elsewhere, and will not be published elsewhere without the permission of the Editor. All the authors have contributed directly to the planning, execution or analysis of the work reported or to the writing of the paper.

Acknowledgments

This research is supported in part by US Army Medical Research and Material Command (USAMRMC) under grant of #W81XWH-11-1-0335 (CUNY RF # 47204-00-01). The authors acknowledge the help of Ms. L. A. Sordillo for helping ordered and kept four normal and cancerous breast tissue specimens for the measurements. We are also grateful to Mr. C. Y. Zhang, Mr. M. Z. Fan, and Dr. X. H. Ni for their assistance in all the experiments and critical suggestions concerning the manuscript.

References


33. Spiro, T. G., Strekas, T. C., Resonance Raman spectra of hemoglobin and cytochrome c: inverse polarization and vibronic scattering. Proc Nat Acad USA 69(9), 2626-2626 (1972).


Received: May 18, 2012; Revised: December 15, 2012; Accepted: December 19, 2012