

Journal of Biomedical Optics

SPIEDigitalLibrary.org/jbo

Stokes shift spectroscopic analysis of multifluorophores for human cancer detection in breast and prostate tissues

Yang Pu
Wubao Wang
Yuanlong Yang
Robert R. Alfano

Stokes shift spectroscopic analysis of multifluorophores for human cancer detection in breast and prostate tissues

Yang Pu, Wubao Wang, Yuanlong Yang, and Robert R. Alfano

City College of the City University of New York, Institute for Ultrafast Spectroscopy and Lasers, Departments of Electric Engineering and Physics, 160 Convent Avenue, New York, New York 10031

Abstract. Stokes shift spectroscopy (S3) offers a novel and simpler way to rapidly recognize spectral fingerprints of multiple fluorophores in complex media such as in tissue. This spectroscopic technique can be used as an effective approach to detect cancer in tissue. The alterations of the measured S3 spectra between cancerous and normal tissues were observed in human breast and prostate samples. In order to obtain the optimal Stokes shift interval, $\Delta\lambda_i$, for the purpose of breast/prostate cancer detection using S3, the S3 spectra of a mixed aqueous solution of tryptophan, nicotinamide adenine dinucleotide, and flavin were measured with different $\Delta\lambda_i$ values. The experimental results analyzed using nonnegative least square method show that there is a reduced contribution from collagen and an increased contribution from tryptophan to the S3 signal of the cancerous tissue as compared with those of the normal tissue. This study indicates that the changes of relative contents of tryptophan and collagen in tissue shown by the S3 spectra may present potential native biomarkers for breast and prostate cancer detection. S3 has the potential to be a new armamentarium. © The Authors. Published by SPIE under a Creative Commons Attribution 3.0 Unported License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: [10.1117/1.JBO.18.1.017005](https://doi.org/10.1117/1.JBO.18.1.017005)]

Keywords: Stokes shift spectroscopy; cancerous breast and prostate tissue; optimal Stokes shift interval; optical biopsy; tryptophan; collagen; nicotinamide adenine dinucleotide.

Paper 12606 received Sep. 12, 2012; revised manuscript received Nov. 13, 2012; accepted for publication Dec. 11, 2012; published online Jan. 7, 2013.

1 Introduction

Optical spectroscopy, in particular the Stokes shift spectroscopy (S3), has the potential to be used as a noninvasive or less invasive technique for cancer detection over other conventional diagnostic methods. Additional advantages of S3 include less time consumption, reproducibility, and minimal invasiveness without removing tissue.¹ S3 falls within the field of “optical biopsy,” a spectroscopic technique to diagnose disease without removing tissue sample, heading toward new advanced tools for medical armamentarium.

Human tissue is mainly composed of epithelial cells, proteins, fat, water, and extracellular matrix of collagen fiber. A number of key fluorophores tryptophan, collagen, elastin, reduced nicotinamide adenine dinucleotide (NADH), and flavin were investigated as the fingerprint molecules in biomedical optics.^{2,3} Tryptophan is an amino acid required by all forms of life for protein synthesis and other important metabolic functions,⁴ and accounted to have major contribution to protein fluorescence. NADH and flavin adenine dinucleotide (FAD) are involved in the oxidation of fuel molecules and can be used to probe changes in cellular metabolism.⁵

Spectral analysis holds a great promise as a tool for diagnosing early stage of carcinomas.^{1-3,6,7} Stokes shift spectroscopy offers a novel and simpler way to rapidly recognize spectral fingerprints of fluorophores in complex mixtures such as in tissue.^{3,8} S3 measurements can be used to acquire enough information of different key fluorophores from one spectrum;

therefore, it can be employed for using one single scan to obtain the most critical fingerprints of main fluorophores which are valuable for cancer detection.⁸ Most recently, the S3 technique has received increasing interest for diagnostics of different types of cancers in human tissues.^{8,9}

The aim of this paper is to show the usefulness of the S3 technique to distinguish the malignant tissues from the normal for the human breast and prostate tissues samples. The optimal Stokes shift interval, $\Delta\lambda_i$, for S3 measurements was investigated and determined. The S3 spectra of the key fluorophores in human prostate and breast tissues were measured and analyzed using nonnegative least square (NNLS) method.¹⁰ The underlying physics and biological basis for the S3 approach were discussed. The linear discriminant analysis (LDA) was used to convert alterations observed in the S3 spectra into valuable information¹¹ for diagnosing the cancerous tissue. Subsequently, the receiver operating characteristic (ROC) curve¹¹ was generated to further evaluate the performance of the NNLS algorithm combined with LDA for diagnosis of human breast/prostate cancer. The results show that S3 spectra can be used as an alternative tool for the detection of breast/prostate cancer to monitor the changes of main fluorophores in tissue.

2 Methods and Samples

For typical organic molecules, the peaks of absorption and emission occur at different wavelengths. The difference between the emission and absorption peaks is known as the Stokes shift depending on emitting molecules and the polarity of the surrounding host environment.⁶ A spectroscopic method combining fluorescence and absorption was proposed to acquire spectra by synchronously scanning the excitation and detection wavelengths with a constant shift interval, $\Delta\lambda_i$, between the excitation and detection wavelengths.^{6,7} This method was named as Stokes

Address all correspondence to: Yang Pu, City College of the City University of New York, Institute for Ultrafast Spectroscopy and Lasers, Departments of Electric Engineering and Physics, 160 Convent Avenue, New York, New York 10031. Tel: +2126505541; Fax: 2126505530; E-mail: puyang@sci.cny.cuny.edu

shift spectroscopy.⁶ While conventional fluorescence spectroscopy uses either a fixed excitation wavelength, λ_{ex} , to produce an emission spectrum or a fixed emission wavelength, λ_{em} , to record an excitation spectrum, the S3 signal is recorded when both λ_{em} and λ_{ex} are simultaneously scanned in the synchronized changing mode of a spectrometer, such as Perkin-Elmer LS 50. In our study, the excitation light with 5 nm spectral width was focused on samples with a spatial size of $\sim 3 \times 1$ mm. The power of incident light was $\sim 0.5 \mu\text{W}$. The scan speed was 250 nm per minute. The fluorescence was collected with a spectral resolution of 0.5 nm.

To investigate the optimal wavelength shift interval of $\Delta\lambda_i$ for cancer detection, the S3 spectra of the mixture of three kinds of key fluorophores, tryptophan, reduced NADH, and flavin, in solution were measured with various $\Delta\lambda_i$, from 20 to 120 nm. The S3 spectra of individual tryptophan, NADH, and flavin in solution, and collagen in solid suspension at were also measured with $\Delta\lambda_i = 40$ nm and used for NNLS analysis for the S3 spectra of the breast and prostate tissue. All fluorophores were obtained commercially from Mallinckrodt Baker, INC.

Human breast and prostate tissue specimens were obtained from the Co-operation Human Tissue Network (CHTN) and the National Disease Research Interchange (NDRI) under an Institutional Review Board (IRB) approval at the City College of New York. CHTN and NDRI provided pathology reports for the tissue specimens. The samples were received on dry ice and were neither chemically treated nor frozen before spectroscopic measurements. The time elapsed between tissue resection and the S3 measurements varied from sample to sample due to different sample sources, but were not exceeding 30 h. When taking the S3 measurements, the cancerous tissue samples were carefully checked to determine the hard parts to locate the cancerous regions before making any measurements. The assessment using hardness is acknowledged to be a simple way to locate regions of malignancy.¹²

3 Experimental Results and Analysis

The S3 spectra of five pairs of cancerous (solid line) and normal (dash line) prostate tissues were measured by setting $\Delta\lambda_i = 40$ nm, and their averaging profile is shown in Fig. 1(a). Figure 1(b) exhibits the averaging S3 spectra of 19 pairs of cancerous (solid line) and normal (dash line) breast tissues acquired with $\Delta\lambda_i = 40$ nm. Each spectral profile was normalized to unit value of 1 (i.e., the sum of squares of the elements in each S3 spectrum was set as 1) before taking average and calculation. The salient differences of the S3 spectra between cancerous and normal tissues displayed in Fig. 1(a) and 1(b) are that $I_c > I_n$ at ~ 294 or ~ 295 nm while $I_c < I_n$ at ~ 340 nm, where I_c and I_n are the S3 spectral intensities of the cancerous and normal tissues, respectively. Another obvious difference between cancerous and normal prostate tissues revealed by the S3 spectra shown in Fig. 1(a) is that the peak intensity at I_{294} is higher than that at I_{340} for the cancerous prostate tissue while this property is inverted in the normal prostate tissue. In addition, there is a small but noticeable spectral difference for the cancerous and normal prostate tissues. The higher intensity at shoulder peak of ~ 394 nm was observed in cancerous prostate tissues in comparison with that in normal prostate tissues. This tiny shoulder peak can also be seen at 385 nm for the cancerous breast tissue by enlarging the spectral profile in the range from 365 to 420 nm as shown in the insert of Fig. 1(b). The S3 spectral differences may reflect the change of tissue structures during the

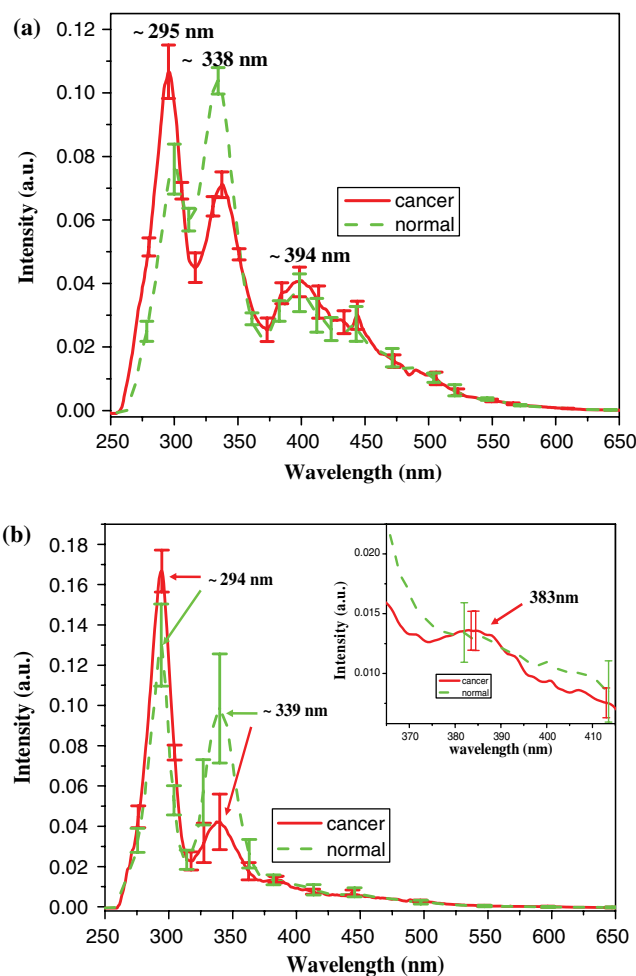


Fig. 1 Average Stokes shift spectra of cancerous (solid) and normal (dash) human (a) prostate and (b) breast tissue samples obtained with the selective $\Delta\lambda_i = 40$ nm. The enlarged spectral profile of breast samples in the wavelength range from 365 to 420 nm is shown as the insert of (b).

development of cancer. In order to understand which biochemical components mainly contribute to these changes, the S3 spectra of the main fluorophores in breast tissue, e.g., tryptophan, collagen, NADH, and flavin, were measured.

Figure 2(a) shows the measured S3 spectrum for each of the individual fluorophore. The spectrum of collagen in solid suspension model was acquired with $\Delta\lambda_i = 40$ nm since the collagen is not soluble in water. The collagen sample was shaken evenly before the measurements. The S3 spectrum of individual tryptophan, NADH, and flavin was measured in the aqueous solution. The S3 spectra of Fig. 2(a) can be used to determine which biochemical components mainly contribute to the S3 spectra of the tissue samples. On comparing Fig. 2(a) with Fig. 1(a) and 1(b), it can be seen that the main peak at ~ 290 nm for the tissue spectra is contributed from tryptophan. The small peak at ~ 340 nm corresponds to collagen, and the very tiny peak at ~ 380 or ~ 390 nm stands for NADH. No obvious peak for flavin was observed. Therefore, it may be concluded that the S3 spectra of tissue samples obtained with $\Delta\lambda_i = 40$ nm is mainly contributed from tryptophan, collagen, and NADH.

In order to investigate the effect of $\Delta\lambda_i$ to the S3 spectra, a mixed aqueous solution of tryptophan, NADH, and flavin with a concentration of $\sim 0.4 \text{ mg/cm}^3$ was measured with different

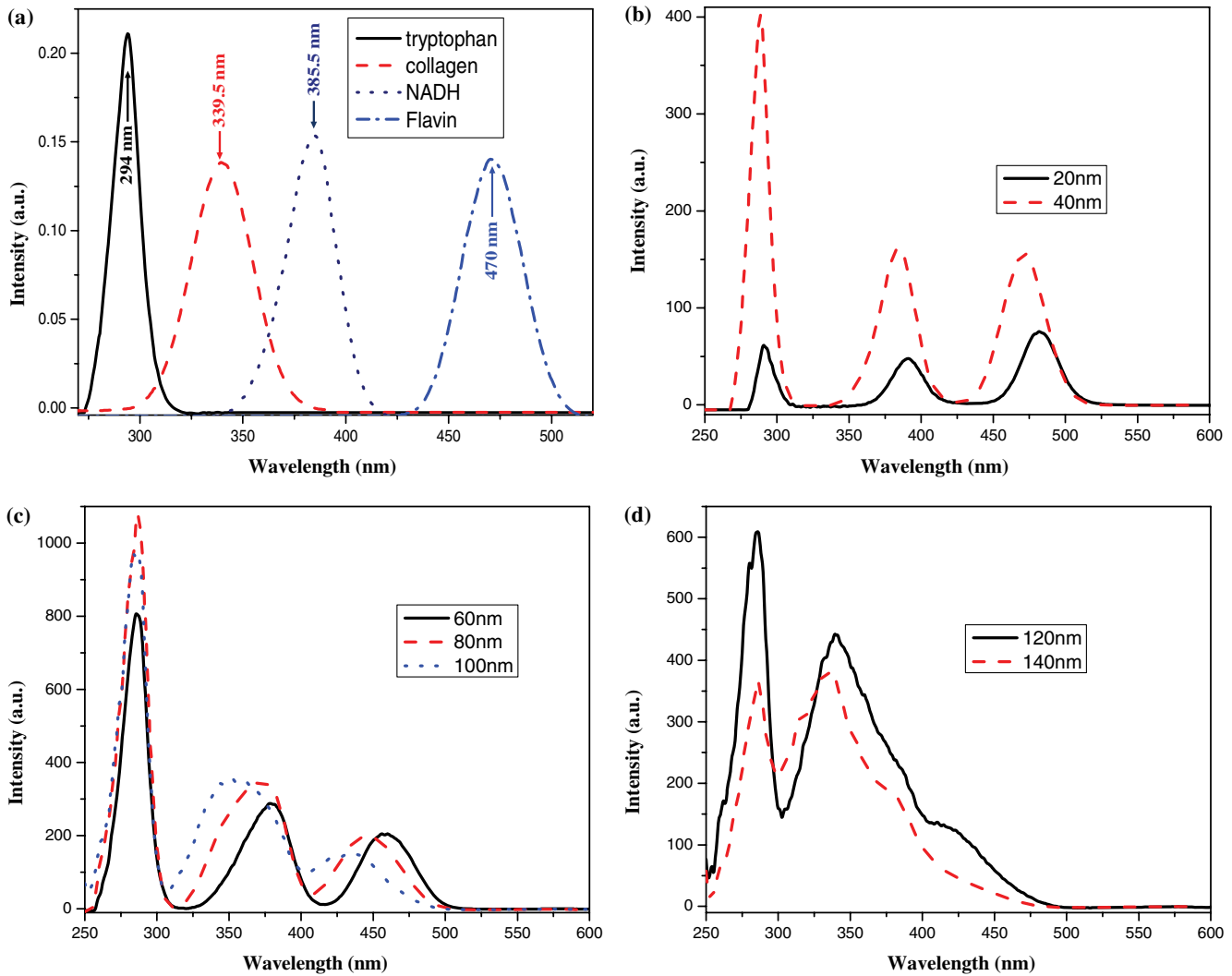


Fig. 2 (a) Stokes shift spectra of individual fluorophore: tryptophan (solid), NADH (dot), and flavin (dash-dot) in solution, and collagen (dash) in solid suspension, and Stokes shift spectra of a mixture solution containing tryptophan, NADH, and flavin obtained with (b) $\Delta\lambda_i = 20$ nm (dot line) and 40 nm (dash line); (c) $\Delta\lambda_i = 60$ nm (solid line), 80 nm (dash line), and 100 nm (dot line); and (d) $\Delta\lambda_i = 120$ nm (solid line) and 140 nm (dash line).

$\Delta\lambda_i$ values. The S3 spectra of the mixture solution obtained with $\Delta\lambda_i = 20$ and 40 nm are displayed as solid and dash lines, respectively, in Fig. 2(b). Figure 2(c) shows the S3 spectra of same mixture solution obtained with $\Delta\lambda_i = 60$, 80, and 100 nm as solid, dash, and dot lines, respectively. The spectra obtained with $\Delta\lambda_i = 120$ and 140 nm are exhibited as solid and dash lines, respectively, in Fig. 2(d). All curves are acquired under same experimental conditions except using different $\Delta\lambda_i$. One salient feature shown in Fig. 2(b) to 2(d) is that with the increase of $\Delta\lambda_i$, the peak intensities contributed from three fluorophores ascend, but drop down at different values of $\Delta\lambda_i$. Another notable feature shown in Fig. 2(b) to 2(d) is that the full width at half maximum (FWHM) of the S3 spectral profiles for all three fluorophores expands monotonously with the increase of $\Delta\lambda_i$. These two salient features shown in Fig. 2(b) to 2(d) are valuable for one to choose the optimal wavelength shift interval $\Delta\lambda_i$, for the S3 spectral study to determine the changes of multifluorophores in tissue.

The FWHM of the S3 spectrum of an individual fluorophore actually indicates the resolution of the spectral signal. A smaller FWHM corresponds to a better spectral resolution.

In fact, a FWHM-based spectral-resolving power (FSRP) can be defined as:

$$P = \frac{\lambda_{\text{range}}}{\Delta\lambda_{\text{FWHM}}}, \quad (1)$$

where λ_{range} is the whole wavelength range for the measurements with effective signal, which can be taken as 500 nm in our study, and $\Delta\lambda_{\text{FWHM}}$ is the FWHM for the S3 spectrum of each fluorophore, which can be obtained from Fig. 2(b) to 2(d). Using Eq. (1) and the values of FWHM and peak intensities shown in Fig. 2(b) to 2(d), the FWHM-based spectral-resolving power and the spectral peak intensities were calculated as a function of $\Delta\lambda_i$ for three fluorophores of interest, and the results are shown in Fig. 3(a) and 3(b), respectively.

Using Fig. 3, one can straightforwardly choose an optimal wavelength shift constant $\Delta\lambda_i$ to obtain the S3 spectra of interest. For signal processing, two most important parameters determine the quality of signals: resolution and magnitude. Figure 3(a) shows that the FSRPs of three fluorophores, tryptophan (squares with a solid line), NADH (circles with a dash line), and flavin (hexagons with a dot line), decrease monotonously with the increase of $\Delta\lambda_i$, indicating that the smaller wavelength

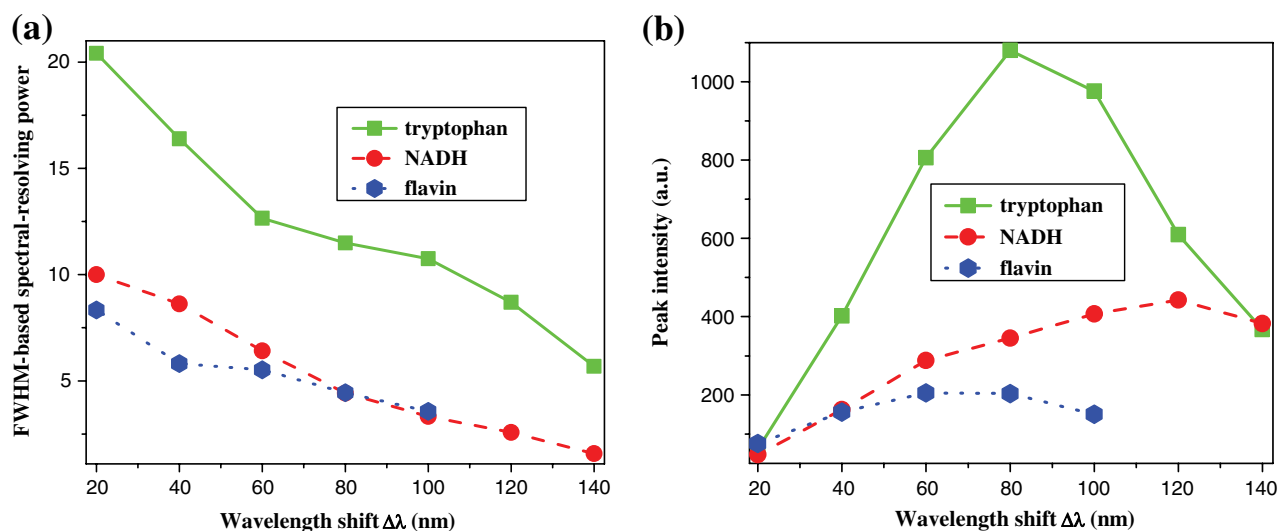


Fig. 3 (a) The FWHM-based spectral-resolving power of the S3 spectra of tryptophan (square with solid line), NADH (circle with dash line), and flavin (hexagon with dot line), as a function of $\Delta\lambda_i$; (b) the peak intensities of Stokes shift spectra of tryptophan (square with solid line), NADH (circle with dash line), and flavin (hexagon with dot line) as a function of $\Delta\lambda_i$.

constant $\Delta\lambda_i$ corresponds to the higher FSRP of spectral signal. Figure 3(b) exhibits: (1) when $\Delta\lambda_i = 20$ nm is used, all of the three fluorophores have approximately same peak but weak intensities; (2) as $\Delta\lambda_i$ increases, the peak intensities of three fluorophores ascend up at first, but descend down at different critical values of $\Delta\lambda_i$. The curve for tryptophan (squares with a solid line) falls at $\Delta\lambda_i = 80$ nm, and that for flavin (hexagons with a dot line) drops at $\Delta\lambda_i = 60$ nm and that for NADH (circles with a dash line) decreases at $\Delta\lambda_i = 120$ nm. These may expose the underlying physical and biological basis for the Stokes shift spectroscopy techniques.

One may recognize that the S3 spectra actually acquire the signal of fluorescence, and the fluorescence intensity is closely related to the quantum yield (QY). The value of $\Delta\lambda_i$ used in the S3 measurements for each biomolecule is determined by the difference of its corresponding peak wavelengths of absorption and emission. When wavelength shift constant $\Delta\lambda_i$ used in the S3 measurements approaches the Stokes shift interval, $\Delta\lambda_{ss}$, defined as the difference of peaks wavelengths between λ_{abs} and λ_{em} , the magnitudes of the S3 spectral signal will access the maximum. When $\Delta\lambda_i$ abandons $\Delta\lambda_{ss}$, the intensity of the S3 spectra will reverse back. To understand Fig. 3(b), the S3-related parameters such as λ_{abs} , λ_{em} , $\Delta\lambda_{ss}$, and QY for main molecules of our interest in tissue samples are listed in Table 1.

When S3 spectra were acquired by setting $\Delta\lambda_i = 20$ nm, the excitation of all three fluorophores is far from their Stokes shift interval, $\Delta\lambda_{ss}$, resulting the smallest signal intensity. With

increase of $\Delta\lambda_i$, the magnitude of signal from tryptophan is boosted drastically because of its higher quantum yield in comparison with NADH and flavin. Since $\Delta\lambda_{ss}$ values of flavin and tryptophan are ~ 70 nm and ~ 70 to 80 nm, respectively, their S3 spectral intensities drops back at $\Delta\lambda_i = \sim 60$ to 80 nm after it exceeds $\Delta\lambda_{ss}$. Similar reason causes peak intensity of NADH regressing back at $\Delta\lambda_i = 120$ nm.

Tryptophan, collagen, NADH, and flavin are key molecules in cancer diagnose using spectroscopy.²⁻⁴ For breast cancer, the most common grading system used in the United States is the Scarff-Bloom-Richardson (SBR) system,¹⁷ which is a breast cancer staging system that examines the cells and tissue structures of tumors to determine how aggressive and invasive the cancer is depending on three features: (1) The percent of the tumor makes the tissue structures change. In cancer, the tissue structures usually become less orderly; (2) the numbers of mitotic figures (dividing cells) observed in a certain magnitude microscope field. One of the hallmarks of cancer is that cells divide uncontrollably; and (3) the nonuniformity of the cell nuclei. The cancerous cells have larger, irregular, and darker cell nuclei than that of normal breast duct epithelial cells.¹⁷ Each of these features is assigned a score ranging from 1 to 3. The lowest score 3 (1 + 1 + 1) is given to well-differentiated tumors with best prognosis while the highest score 9 (3 + 3 + 3) for poorly differentiated tumor is the worst prognosis.¹⁷ According to the features described by the SBR system, higher cell density, uncontrollable cells-dividing, and nonuniform larger cellular

Table 1 S3-related values and quantum yield of key biomolecules.

Molecule	λ_{abs} (nm)	λ_{em} (nm)	$\Delta\lambda_{ss}$ (nm)	Quantum yield
Tryptophan	270 to 280	350	70 to 80	0.3 to 0.35 (Ref. 13)
Collagen	340	380	40	0.3 (Ref. 14)
NADH	340	440 to 460	100 to 120	0.019 to 0.05 (Refs. 15 and 16)
Flavin	450	520	70	0.03 (Ref. 13)

nuclei are characteristics for cancerous breast cells; therefore the changes of fluorescence from the main fluorophores inside cells (e.g., tryptophan, NADH, and flavin) should be expected. The primary fluorophore in the breast tissue extracellular matrix is type I collagen.¹⁸ For invasion and subsequent metastasis, tumor cells degrade the surrounding extracellular matrix, which is composed mainly of type I collagen.¹⁸ Understanding these changes during breast cancer evolution is critical to reveal the contributions of the biochemical components in tissue for the spectroscopic features.

It can be seen from Fig. 3(a) and 3(b) that if spectral resolution is the only consideration, optimal $\Delta\lambda_i$ should be chosen as small as possible. On the other hand, in order to enhance the signal-to-noise ratio (SNR), the optimal $\Delta\lambda_i$ should be chosen as close as possible to $\Delta\lambda_{ss}$ by means of acquiring the highest signal magnitude of the fluorophore of interest. In the application of breast cancer detection, the alteration of biochemical components due to the evolution of tumor allows an optimal and easy way to highlight the difference between cancerous and normal tissues using S3 techniques with selective optimal $\Delta\lambda_i$. In the spectral analysis for cancer detection, the differences of emission of a fluorophore among different tissue specimens and patients pronounces unreliability to compare the absolute emission intensities of certain biomarkers for different samples. Although fluorescence intensity is proportional to the number of fluorophores and the contents of the fluorophores can be calculated from the measured intensity, the calculation of the absolute contents requires complicated calibration algorithm and the information of the tissue environment such as pH-values, polarity, temperature, and viscosity. These make impossible to compute the absolute contents of the fluorophores using spectral method. An easy way is to find and use an unchanged component as a reference to measure the changes of the fluorophores of our interest. It is better to use the changes of relative contents of tryptophan and collagen in the case of breast cancer detection because of evidence of increase of tryptophan and the decrease of collagen in cancerous tissue.^{17,18} Since $\Delta\lambda_{ss} = 40$ to 50 nm for collagen and $\Delta\lambda_{ss} = 70$ to 80 nm for tryptophan, $\Delta\lambda_i = 40$ nm should be chosen as optimal scan wavelength interval, which gives balance between the resolution and the SNR. Furthermore, the S3 spectra with $\Delta\lambda_i = 40$ nm can be used to investigate the signal arisen from tryptophan and collagen in tissue, which vary inversely in cancerous and normal tissues. Therefore, the S3 with selective $\Delta\lambda_i = 40$ nm can highlight the difference between cancerous and normal tissues with the inverse spectral property exhibited in Fig. 1.

To investigate changes of the relative contents of fluorophores in tissue from the S3 spectra, a forward analyzing method, namely NNLS,¹⁰ was applied to extract relative contents of the fluorophores, e.g., tryptophan, collagen, and NADH, using their individual S3 spectrum obtained with $\Delta\lambda_i = 40$ nm shown in Fig. 2(a) and the S3 spectra of the cancerous and normal breast tissues as shown in Fig. 1(b). Figure 4(a) shows the NNLS-extracted relative contents of tryptophan versus collagen, Fig. 4(b) displays the extracted relative contents of collagen versus NADH, and Fig. 4(c) exhibits the extracted relative contents of tryptophan versus NADH for different cancerous (squares) and normal (circles) breast tissues. The most salient feature of Fig. 4(a) is that all data points for the normal tissues locate in the left-upper of the data for the cancerous tissues, indicating that the relative contents of collagen in normal tissues are higher in comparison with the cancerous tissues while the relative

contents of tryptophan in normal tissues are lower than those in the cancerous tissues. Figure 4(b) shows that the relative contents of collagen are lower, but the relative contents of NADH are higher in cancerous tissue in comparison with the normal tissues. Figure 4(c) provides reproducible evidence of increase of the relative contents of tryptophan and NADH in cancerous breast tissues.

Figure 4(a) to 4(c) obviously shows that the relative content of collagen in cancerous breast tissue is lower than that in normal breast tissue, but the relative contents of tryptophan and NADH are larger in cancerous breast tissues in comparison with that in normal breast tissue. This observation is in good agreement with other group's studies performed for cancerous and normal breast cells and tissues.^{3,4,17-21}

In a diagnosis test for cancer, the test outcome can be positive (cancer) or negative (healthy). LDA is a method to find a linear combination of features which separates two or more classes of objects or events,¹¹ which is very useful for the two-group classification. To evaluate the potential of a diagnosis method, the following statistic terms are usually used: (1) true positive: defining as a cancerous sample correctly diagnosed as malignant; (2) false positive: defining as a healthy sample incorrectly identified as malignant; (3) true negative: defining as a healthy sample correctly identified as healthy; and (4) false negative: defining as a cancerous sample incorrectly identified as healthy.

The sensitivity and specificity then can be calculated by:

$$\text{Sensitivity} = \frac{\# \text{ of True Positive}}{\# \text{ of True Positive} + \# \text{ of False Negative}}$$

$$\text{Specificity} = \frac{\# \text{ of True Negative}}{\# \text{ of True Negative} + \# \text{ of False Positive}}$$

The criteria of categorizing the true or false positive and negative groups in our study were determined by LDA model. The separating lines on the scatter plots shown in Fig. 4(a) to 4(c) were loaded by the LDA model for the three diagnostically significant fluorophores, e.g., tryptophan, collagen, and NADH. To evaluate the S3 spectra analysis combined with LDA as a criterion for effective diagnostic algorithms for breast tissue classification, the sensitivity and specificity for our diagnostic results were calculated and are summarized in Table 2.

The performance of a two-group classification problem is typically evaluated with an ROC curve, which is a graphical plot of true positive rate versus false positive rate.¹¹ Accuracy is measured by the area under the ROC curve (AUC). A rough guide for classifying the accuracy of a diagnostic test is the traditional academic point system: (A) excellent = 0.90 to 1; (B) good = 0.80 to 0.90; (C) fair = 0.70 to 0.80; (D) poor = 0.60 to 0.70; and (F) fail = 0.50 to 0.60. The ROC curves shown in Fig. 4(d), 4(e), and 4(f) were generated from the scatter plot shown in Fig. 4(a), 4(b), and 4(c) to determine the correct or incorrect classification of cancerous and normal breast tissues using different pairs of diagnostically significant fluorophores, respectively. Figure 4(d) shows the ROC curve generated from the pairs of diagnostically significant fluorophores of tryptophan versus collagen, Fig. 4(e) displays the ROC curve generated from the pairs of diagnostically significant fluorophores of collagen versus NADH, and Fig. 4(f) exhibits ROC from tryptophan versus NADH. The AUC calculated from the ROC shown in Fig. 4(d), 4(e), and 4(f) is 0.956, 0.97, and 0.964,

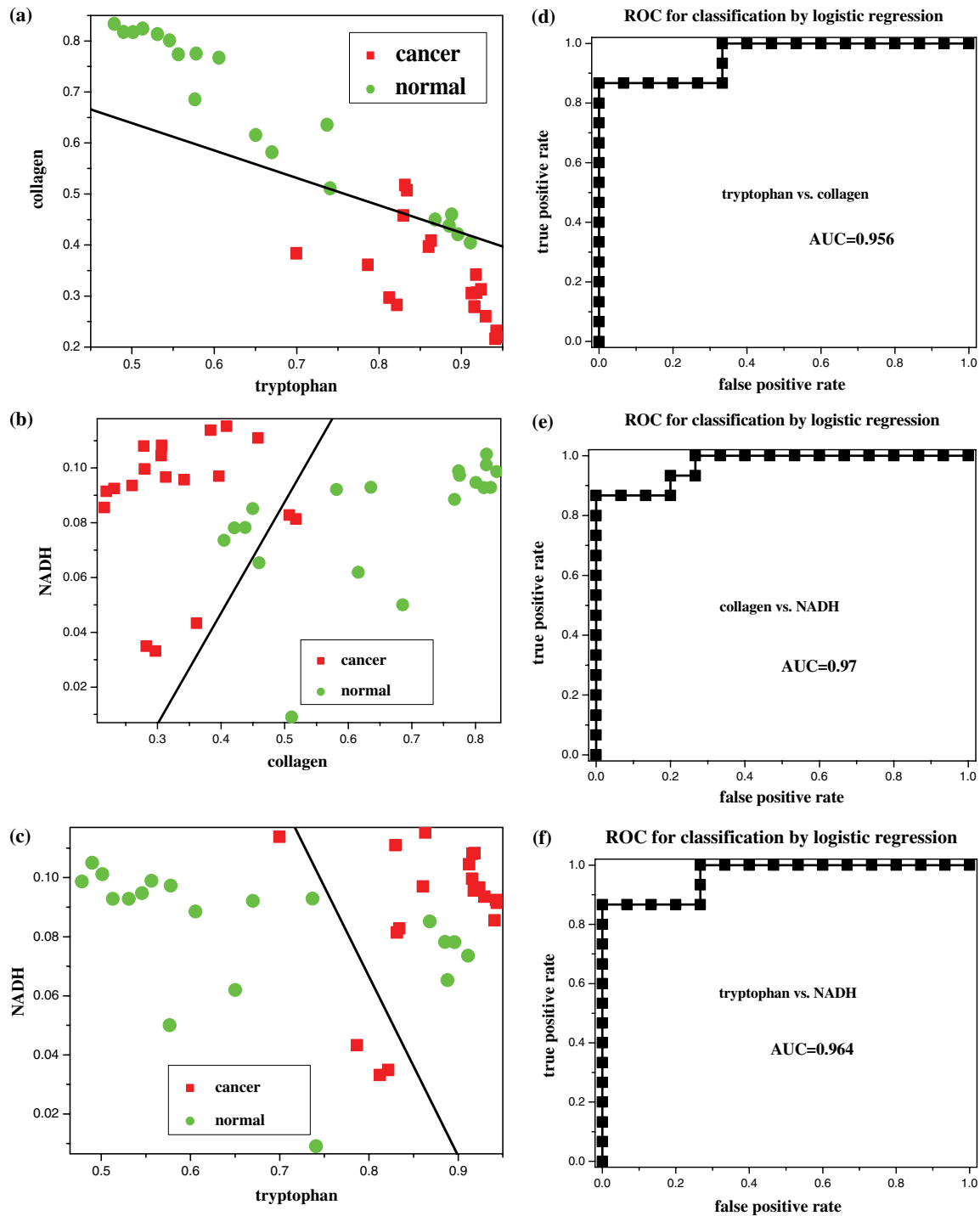


Fig. 4 Scatter plots of the relative contents of (a) tryptophan versus collagen, (b) collagen versus NADH, and (c) tryptophan versus NADH by analyzing Stokes shift spectra recorded with the selective $\Delta\lambda_i = 40$ nm. The separating lines were calculated using the LDA method. The calculated ROC corresponding to the relative contents of (d) tryptophan versus collagen; (e) collagen versus NADH; and (f) tryptophan versus NADH as the pairs of diagnostically significant fluorophores to evaluate performance of classifying breast tissues into cancerous and normal groups based on the S3 spectra and NNLS analysis.

respectively, demonstrating the excellent efficacy of the S3 spectra with selective $\Delta\lambda_i = 40$ nm and the NNLS analysis combined with LDA as a promising diagnostic tool for breast cancer.

4 Discussion

These initial results explore the efficacy of the S3 spectra combined with NNLS and LDA methods for cancer detection in

human breast and prostate tissue samples. It was also found that tryptophan, collagen, and NADH can be used as potential biomarkers for cancer detection using the S3 techniques with the selective optimal $\Delta\lambda_i = 40$ nm.

The conventional spectral approaches applied in tissue optics studies are absorption or optical density (OD), fluorescence (emission and/or excitation), and excitation-emission matrix

Table 2 The sensitivity and specificity of the S3, NNLS, and LDA analysis for breast cancer detection.

Evaluated components	Sensitivity	Specificity
Tryptophan versus collagen	84%	74%
Collagen versus NADH	90%	79%
Tryptophan versus NADH	79%	74%

(EEM) measurements. In the tissue absorption measurements, only few chromophores can match the detectable level. The fluorescence spectroscopy including either emission or excitation can detect low-level fluorophores, but because of the fixed pump or detecting wavelength, the strongest emission or excitation spectral signal can be acquired only for one or two fluorophores. In addition, the emission signals of most fluorophores have very wide FWHM. Table 3 is the comparison of FWHMs of the S3 spectra obtained with $\Delta\lambda_i = 40$ nm and conventional emission spectra for the four main fluorophores of our interest: tryptophan, collagen, NADH, and flavin.^{3,6,22}

It can be seen from Table 3 that the conventional emission spectral measurements provide poor resolution and much less information than the S3.

Although EEM can be used to ensure the coverage of all endogenous fluorophores, the data acquisition is extremely time consuming and, thus, not suitable for clinic use. Furthermore, redundant information of EEM conceals alteration of the spectral fingerprints of cancerous and benign breast tissues. In contrast, S3 spectral measurements can be used to acquire enough information of multiple key fluorophores at much low content and reach relative higher resolution with a single scan (as compared with EEM). The S3 approach can dramatically reduce data acquisition time and keep the reasonably high classification accuracy. The most obvious advantage of S3 over other conventional spectral methods is that it gives a simple and efficient way to highlight differences between normal and diseased tissues, which is more robust, evident, and pronounced.

The S3 spectra for cancer detection have also been investigated for different types of tissues other than breast. Reproducible results were observed. For example, similar changes of biomarker for cancerous and normal prostate tissue were observed as shown in Fig. 1(a), where higher peak intensities of the S3 spectra at ~ 290 nm for tryptophan and ~ 390 nm for NADH, and lower peak at ~ 340 nm for collagen in the cancerous prostate tissues, indicating that there is a reduced contribution from the emission of collagen and increased contributions from tryptophan and NADH in cancerous prostate tissues as compared with normal tissue. These reproducible changes

Table 3 Comparison of FWHM between the S3 spectra obtained with $\Delta\lambda_i = 40$ nm and conventional emission spectra.

Molecules	Tryptophan	Collagen	NADH	Flavin
FWHM of S3 ($\Delta\lambda_i = 40$ nm)	16	32	29	43
FWHM of emission	68 (Ref. 6)	64 (Ref. 3)	93 (Ref. 3)	58 (Ref. 22)

were also observed in lung cancer tissues in limited number of samples. Further investigation needs to be achieved in large number of different types of cancerous tissue specimens.

The histological imaging analysis is needed to confirm our spectral study for diagnosis purpose of breast/prostate cancer in collaboration with pathology experts. The stained normal and cancerous breast/prostate tissue histological slides corresponding to the pairs of the cancerous and normal tissues used for the spectral measurements will be needed for microscopy study to investigate the presence of these spectral biomarkers, tryptophan, collagen, and NADH, as a function of types and stages of cancer, specifically SBR scores of breast cancer, which is used to determine the aggressiveness and invasiveness of the breast cancer, and Gleason grades of prostate cancers, which is used to evaluate the prognosis and aggressiveness of men with prostate cancer.

5 Conclusion

Stokes shift spectroscopy method offers a robust and efficient way to rapidly measure spectral fingerprints of fluorophores in complex media such as in tissue and highlights the difference between cancerous and normal tissues. This paper explicitly discloses the underlying basis for how and why this spectral method is superior in comparison with absorption and fluorescence spectroscopic methods. It also specifies how to select the optimal $\Delta\lambda_i$ for breast cancer detection. This study demonstrates that the S3 method can be used to acquire information for different key fluorophores from one S3 spectrum and to investigate the changes of the relative contents of the key fluorophores in breast/prostate tissues during the development of cancer. The great advantage of S3 is that a much smaller number of data points are required while much more fluorophores can be detected. S3 can be a new optical tool for medical armamentarium.

Acknowledgments

This research is supported in part by U.S. Army Medical Research and Materiel Command (USAMRMC) under grant of # W81XWH-11-1-0335 (CUNY RF # 47204-00-1) and W81XWH-08-1-0717 (CUNY RF # 47170-00-01). The authors acknowledge the help of CHTN and NDRI for providing normal and cancerous breast/prostate tissue samples for the measurements.

References

1. R. R. Alfano et al., "Laser induced fluorescence spectroscopy from native cancerous and normal tissue," *IEEE J. Quantum Electron.* **20**(12), 1507–1511 (1984).
2. Y. Pu et al., "Changes of collagen and NADH in human cancerous and normal prostate tissues studied using fluorescence spectroscopy with selective excitation wavelength," *J. Biomed. Opt.* **15**(4), 047008 (2010).
3. Q. Liu, G. Grant, and T. Vo-Dinh, "Investigation of synchronous fluorescence method in multicomponent analysis in tissue," *IEEE J. Quantum Electron.* **16**(4), 927–940 (2010).
4. J. R. Moffett and M. A. Nambodiri, "Review article: tryptophan and the immune response," *Immunol. Cell Biol.* **81**(4), 247–265 (2003).
5. D. K. Bird et al., "Metabolic mapping of MCF10A human breast cells via multiphoton fluorescence lifetime imaging of the coenzyme NADH," *Cancer Res.* **65**(19), 8766–8773 (2005).
6. Y. Pu et al., "Native fluorescence spectroscopic evaluation of chemotherapeutic effects on malignant cells using nonnegative matrix factorization analysis," *Technol. Cancer Res. Treat.* **10**(2), 113–120 (2011).

7. Y. Pu et al., "Time-resolved spectral wing emission kinetics and optical imaging of human cancerous and normal prostate tissues," *Opt. Commun.* **282**(21), 4308–4314 (2009).
8. Y. Pu et al., "Stokes shift spectroscopy highlights differences of cancerous and normal human tissues," *Opt. Lett.* **37**(16), 3360–3362 (2012).
9. J. Ebenezar et al., "Stokes shift spectroscopy pilot study for cancerous and normal prostate tissues," *Appl. Opt.* **51**(16), 3642–3649 (2012).
10. A. De Juan and R. Tauler, "Multivariate curve resolution (MCR) from 2000: progress in concepts and applications," *Crit. Rev. Anal. Chem.* **36**(3–4), 163–176 (2006).
11. L. Parra, "Linear discrimination single trial detection in EEG/MEG—biomedical signal processing and signal modeling," <http://bme.cuny.cuny.edu/faculty/parra/teaching/biomed-dsp/class8.pdf>.
12. A. Shmilovici, "Incomplete tumor volume reduction may improve cancer prognosis," *Med. Hypotheses* **68**(6), 1236–1239 (2007).
13. M. Sauer, J. Hofkens, and J. Enderlein, "Fluorophores and fluorescent labels," Chapter 2 in *Handbook of Fluorescence Spectroscopy and Imaging: From Single Molecules to Ensembles*, pp. 31–60, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany (2011).
14. M.-A. Mycek and B. W. Pogue, Eds., *Handbook of Biomedical Fluorescence*, Marcel Dekker, New York (2003).
15. A. Baici et al., "Fluorescence properties of reduced thionicotinamide—adenine dinucleotide and of its complex with octopine dehydrogenase," *Eur. J. Biochem.* **83**(2), 601–607 (1978).
16. J. R. Lakowicz, *Principles of Fluorescence Spectroscopy*, Kluwer Academic, Dordrecht, The Netherlands (1999).
17. H. J. G. Bloom and W. W. Richardson, "Histological grading and prognosis in breast cancer; a study of 1409 cases of which 359 have been followed for 15 years," *Br. J. Cancer* **11**(3), 359–377 (1957).
18. G. Fenhalls, D. M. Dent, and M. I. Parker, "Breast tumor cell-induced down-regulation of type I collagen mRNA in fibroblasts," *Br. J. Cancer* **81**(7), 1142–1149 (1999).
19. I. Balslev et al., "The Nottingham prognostic index applied to 9,149 patients from the studies of the Danish Breast Cancer Cooperative Group (DBCG)," *Breast Canc. Res. Treat.* **32**(3), 281–290 (1994).
20. J. F. Simpson et al., "Prognostic value of histologic grade and proliferative activity in axillary node-positive breast cancer: results from the Eastern Cooperative Oncology Group Companion Study, EST 4189," *J. Clin. Oncol.* **18**(10), 2059–2069 (2000).
21. A. S. Haka et al., "Diagnosing breast cancer by using Raman spectroscopy," *PNAS* **102**(35), 12371–12376 (2005).
22. Y. Sun et al., "Biomarkers spectral subspace for cancer detection," *J. Biomed. Opt.* **17**(10), 107005 (2012).