Label-free detection of serum proteins using surface-enhanced Raman spectroscopy for colorectal cancer screening

Jing Wang
Duo Lin
Juqiang Lin
Yun Yu
Zufang Huang
Yanping Chen
Jinyong Lin
Shangyuan Feng
Buhong Li
Nenrong Liu
Rong Chen

Label-free detection of serum proteins using surface-enhanced Raman spectroscopy for colorectal cancer screening

Jing Wang,[a,†] Duo Lin,[b,†] Juqiang Lin,[a,*] Yun Yu,[b] Zufang Huang,[a] Yanping Chen,[c] Jinyong Lin,[a] Shangyuan Feng,[a] Buhong Li,[a] Nenrong Liu,[a] and Rong Chen[a,*]
[a]Fujian Normal University, Ministry of Education and Fujian Provincial Key Laboratory for Photonics Technology, Key Laboratory of Optoelectronic Science and Technology for Medicine, Fuzhou 350007, China
[b]Fujian University of Traditional Chinese Medicine, College of Integrated Traditional Chinese and Western Medicine, Fuzhou 350122, China
[c]Teaching Hospital of Fujian Medical University, Fujian Provincial Cancer Hospital, Fuzhou 350014, China

Abstract. Surface-enhanced Raman scattering (SERS) spectra of serum proteins purified from human serum samples were employed to detect colorectal cancer. Acetic acid as a new aggregating agent was introduced to increase the magnitude of the SERS enhancement. High-quality SERS spectra of serum proteins were acquired from 103 cancer patients and 103 healthy volunteers. Tentative assignments of SERS bands reflect that some specific biomolecular contents and protein secondary structures change with colorectal cancer progression. Principal component analysis combined with linear discriminant analysis was used to assess the capability of this approach for identifying colorectal cancer, yielding diagnostic accuracies of 100% (sensitivity: 100%; specificity: 100%) based on albumin SERS spectroscopy and 99.5% (sensitivity: 100%; specificity: 99%) based on globulin SERS spectroscopy, respectively. A partial least squares (PLS) approach was introduced to develop diagnostic models. An albumin PLS model successfully predicted the unidentified subjects with a diagnostic accuracy of 93.5% (sensitivity: 95.6%; specificity: 91.3%) and the globulin PLS model gave a diagnostic accuracy of 93.5% (sensitivity: 91.3%; specificity: 95.6%). These results suggest that serum protein SERS spectroscopy can be a sensitive and clinically powerful means for colorectal cancer detection. © 2014 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: 10.1117/1.JBO.19.8.087003]

Keywords: surface-enhanced Raman scattering; colorectal cancer; serum protein; principal component analysis; linear discriminant analysis; partial least squares approach.

Paper 140199RR received Mar. 26, 2014; revised manuscript received Jul. 26, 2014; accepted for publication Jul. 28, 2014; published online Aug. 19, 2014.

1 Introduction

Colorectal cancer, the third most common cancer and the fourth leading cause of cancer death worldwide, is a major public health problem.[1] In most countries, the incidence rates have increased during the past decades.[2] Detection of cancer at an early stage yields an excellent prognosis, with a 5-year survival rate of over 90%.[3] Thus, an early screening will help to improve the crucial survival rate.

However, colorectal cancer screening is particularly challenging, especially in countries where there are high risks for colorectal cancer.[4] Conventional strategies for the early detection of colorectal cancer include the examination of fecal occult blood test (FOBT), flexible sigmoidoscopy, colonoscopy, double-contrast barium enema, and computed-tomographic colonography (CTC). However, FOBT, flexible sigmoidoscopy, and double-contrast barium enema demonstrate relatively poor sensitivities for colorectal cancer detection (e.g., 30% to 80% with FOBT, 35% to 70% with flexible sigmoidoscopy, and 48% with double-contrast barium enema), and CTC detection suffers from the risk of repeated ionizing radiation for patients.[4] Colonoscopy, though it is the "gold standard" for

colorectal cancer diagnosis, is of high cost and causes pain and discomfort to patients, hampering its application in mass screenings. Therefore, developing a patient-friendly and sensitive method for colorectal cancer screening is imperative.

An optical technique, with the advantage of providing objective and specific information of biochemical changes during cancer development, is being extensively employed for detection and analysis of some diseases.[5–7] Raman spectroscopy, in particular, is additionally attractive as a potential diagnostic tool that can effectively provide information concerning the structures and chemical compositions of biological materials at a molecular level.[1,8] Raman-based detections of serum, circulating tumor cells, and tissue have already been exploited for colorectal cancer detection and diagnosis.[1,9–12] However, due to the weak Raman signal, some of these detections require a very long time for spectral integration. By contrast, surface-enhanced Raman scattering (SERS) technology can easily overcome this issue by significantly enhancing the Raman signal, which expands the applicability of Raman-based analysis in complex biological samples.

Albumin and globulin, which make up most of the total serum proteins, help with the proper functioning of body processes. A marked depression of serum albumin in the presence of various progressive cancers has been well recognized since at

---

*Address all correspondence to: Juqiang Lin, E-mail: jqlin@fjnu.edu.cn; Rong Chen, E-mail: chenr@fjnu.edu.cn

†These authors have made equal contributions as cofirst authors.

least 1950.[13] Serum albumin is generally used to assess the nutritional status, severity of disease, disease progression, and prognosis.[14] The rate of albumin synthesis, which is determined by the supply of amino acids, significantly changes in the critically ill.[15] In addition, the binding of substances (e.g., fatty acid) to albumin can also induce a dramatic conformational change in protein. Albumin bound-lipids have been found to increase in patients due to the enhancement of metabolic pathways required for cancer cell proliferation.[16] Hence, it is possible to monitor cancer progression through analyzing the changes of the secondary structure of albumin. Globulins have multiple functions, depending on their types: gamma globulins (primarily associated with immune system function), beta globulins (primarily associated with hormone transport), alpha-1 globulins, and alpha-2 globulins (primarily associated with clotting function). One group of gamma globulins is the immunoglobulins, which are also known as "antibodies." They have special shapes that recognize, bind to and surround foreign substances, including bacteria or virus, so that scavenger cells can destroy the foreign substances and flush them out of the body. Accordingly, the protein content in the body may provide us with some clinical information regarding a patient's general status.[17]

In 2012, Chen et al. employed the SERS spectrum of circulating ribonucleic acid (RNA) in serum to detect colorectal cancer.[18] However, the SERS spectrum of RNA was difficult to obtain due to the low concentration of RNA in the serum. In 2011, Lin et al. detected colorectal cancer based on the SERS spectrum of the serum.[19] Nonetheless, the SERS characteristics of the serum were most dominated by uric acid, a metabolite whose blood concentration depends on factors such as sex, age, and therapeutic treatments, and were vulnerable to the interference from exogenous substances such as drugs.[20] This suggests that the spectral differences between test and control groups might be masked by greater differences due to the large interindividual variability of uric acid and exogenous substance levels. To further improve these blood test-based noninvasive cancer detection technologies, our research group developed a label-free SERS detection method based on serum protein spectroscopic fingerprints for nasopharyngeal, gastric, and hepatocellular cancer detections.[21–23] This approach has achieved a sensitivity and specificity of 100% for discriminating these three kinds of cancers from the normal group. To date, the potential of serum protein SERS for colorectal cancer detection has not yet been reported in the literature.

In this study, we assess the utility of serum protein-based SERS technology for colorectal cancer detection. This method provides intrinsic fingerprints of proteins for cancer detection by purifying albumin and globulin from complex blood samples, but without using any external labels. One hundred and three colorectal cancer samples and 103 control samples were collected for analysis. Multivariate statistical analyses including principal component analysis (PCA), linear discriminant analysis (LDA), and the partial least squares (PLS) approach were employed to fully and rigorously demonstrate the diagnostic ability of this approach, as well as to construct a diagnostic model to predict the "unknown" samples.

2 Materials and Methods

2.1 Preparation of Human Serum Samples

Two subject groups were involved in this work: the first group consisted of 103 patients with confirmed clinical and histopathological diagnoses of colorectal cancer, and the second group consisted of 103 healthy volunteers as the control group. All participants had similar ethnic and socioeconomic backgrounds and were from the Fujian Provincial Cancer Hospital. The mean age for the cancer group was 55.7 years and for the control group it was 42.3 years. All patients were untreated patients with primary colorectal cancer. They consisted of 31 cases of T1–T2 stage and 72 cases of T3–T4 stage. More detailed clinical information for these patients is given in Table 1. All the participants gave informed consent before sample collection. Whole blood was collected in glass tubes and allowed to clot at room temperature for 15 to 30 min. Then serum samples were obtained by centrifugation (3000 rpm, 2 min).

2.2 Preparation of Protein-Silver Nanoparticles

Similarly to the previous reports, serum proteins were purified from blood serum with membrane electrophoresis (ME).[21–23] Silver nanoparticles (Ag NPs) were reduced by hydroxylamine hydrochloride according to the method reported by Leopold and Lendl.[24] The absorption spectrum of Ag NPs was recorded using a Perkin-Elmer Lambda 950 Spectrophotometer (Waltham, Massachusetts) to characterize the average particle size. Figure 1(a) shows the simplified schematic of the main procedure for obtaining the mixture of serum protein and Ag NPs. Briefly, the blood serum was first blotted onto the cellulose acetate membrane to perform ME. After ME, the membrane was rinsed with a mixed solution of 95% ethanol, glacial acetic acid, and distilled water with a volume ratio of 9∶1∶10 to purify serum proteins from other materials in the blood serum. The membrane containing serum proteins was then equally divided into two parts along a vertical line. Half of the membrane was stained with 0.5% amino black 10B to label the locations of serum proteins (albumin and globulin) for reference, and the serum proteins in the remaining half were cut down according to the labeled positions. Acetic acid was added to dissolve the membrane and Ag NPs were subsequently added and mixed to enhance the protein signal. After 10 min of incubation, acetic acid was again added to the protein-Ag NPs mixture as an

Table 1 Clinical information on colorectal cancer patients and healthy volunteers.

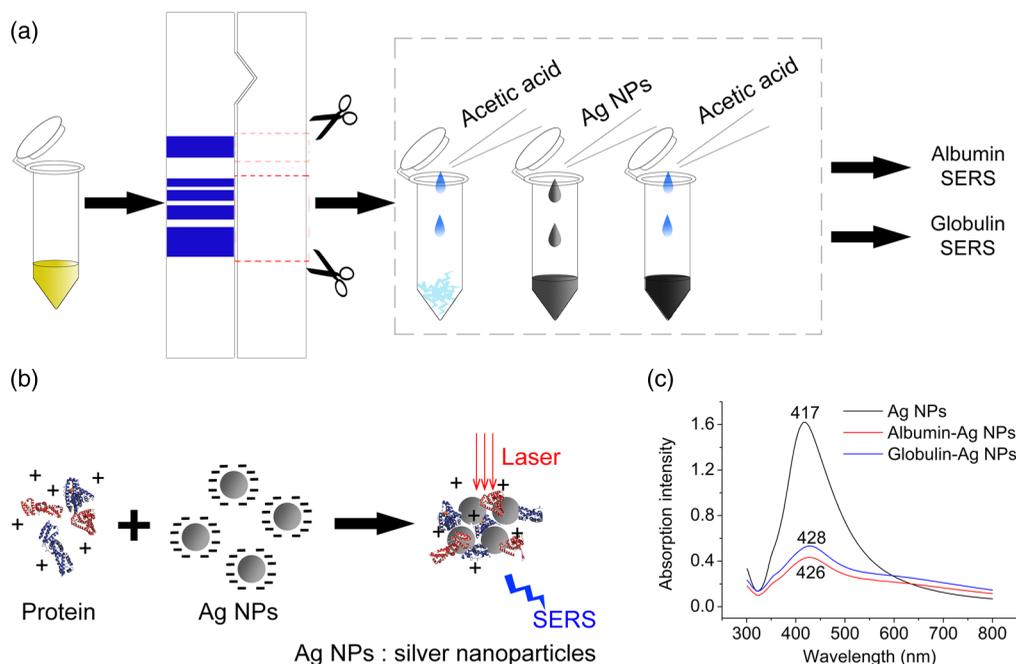|	Colorectal cancer ($n = 103$)	Healthy controls ($n = 103$)
Age		
Mean	55.7	42.3
Median	56	43
Gender		
Male	63	53
Female	40	50
Cancer stage		
T1 to T2	31	N/A
T3 to T4	72	N/A

Fig. 1 (a) Schematic of the preparation procedure of serum protein-Ag NPs mixture. (b) Aggregation protocol of Ag NPs for label-free protein surface-enhanced Raman scattering (SERS) detection. (c) The UV-vis absorption spectra of Ag NPs, albumin-Ag NPs, and globulin-Ag NPs.

aggregating agent to promote the aggregation of protein and Ag NPs, with the aim of further magnifying the protein Raman signal. The mixture was incubated at 37°C and kept continually stirred for 30 min. Prior to the SERS measurement, the supernatant solution (protein-Ag NPs) was deposited on an aluminum sheet and dried at 40°C using a constant-temperature drying oven. Each sample was analyzed with three repeated SERS measurements, then the average was recorded as the final SERS spectrum.

2.3 SERS Spectral Measurement and Preprocess

The SERS spectrum was acquired with a 10 s integration time in the range of 400 to 1800 cm^{-1} using a confocal Raman microspectrometer (inVia System, Renishaw plc, Gloucestershire, United Kingdom). A 785-nm diode laser was focused through a Leica 50× objective to excite the samples. SERS spectra typically resulted from ~0.1 mW. A Peltier cooled charge-coupled device camera (400 × 578 pixels) and the software package of WIRE 2.0 (Renishaw) were employed for spectral acquisition. The 520 cm^{-1} band of a silicon wafer was used for frequency calibration.

Prior to further analysis, the original SERS spectra were baseline-corrected to remove the fluorescence background using the Vancouver Raman Algorithm (a five-order polynomial fitting algorithm) which was developed by BC Cancer Agency and University of British Columbia.[25] Origin eight software (OriginLab Inc., Northampton, Massachusetts) was employed to normalize each of the corrected SERS spectrum by integrating the area under the curve from 400 to 1800 cm^{-1}. Area normalization of the spectroscopic data was performed to compensate for gross differences in the spectral response due to the physical effects rather than the compositional properties of the samples.

2.4 Data Analysis

2.4.1 Linear discriminant analysis

To evaluate the ability of the proposed method in regard to specific discrimination between the normal group and the colorectal cancer group, LDA was performed on the entire dataset (206 average spectra) to identify the directions of the maximum discrimination between groups. However, this has a tendency for over-fitting when the training samples are small compared to the dimensionality. PCA is less sensitive to over-fitting, therefore, a hybrid model which incorporates both LDA and PCA criteria by means of a regularization parameter was proposed. Briefly, PCA was first performed on the spectral data to transform a set of closely correlated variables into uncorrelated ones called PCs. These PCs explain significant differences in the dataset. The optimal number of PCs for LDA was properly determined according to the classification accuracy and cumulative variance.[26] These PCs were then used to develop a classification model for the differentiation of the normal group and the colorectal cancer group. Discriminant scores, obtained from the estimation equation assuming equal *a priori* probabilities of group membership (i.e., independent of differences in the size of the groups), were selected to visualize the classification. PCA-LDA was performed on the normalized spectral data using SPSS 19.0 software package (SPSS Inc., Chicago, Illinois).

2.4.2 Partial least squares regression

To test the predictive power of serum protein SERS spectral data for colorectal cancer detection, the PLS approach was performed on the same spectral data. Latent variables (LVs) were calculated to explain the diagnostic relevant variations rather than the significant differences in the dataset. The use of the PLS approach would be beneficial for spectroscopic

diagnostics since it provides group affinity information (e.g., all samples belong to classes 1 or −1). The optimal number of LVs included in a PLS model and the performance of the PLS model were validated in an unbiased manner using a leave-one-out, cross-validation method. The correlation coefficient (R), R^2, and the root-mean standard error (RMSE) were calculated to assess the fitting of the models. Equations for these parameters can be found in Ref. 27.

The entire dataset (consisting of 206 spectra) was divided into two parts: a training set (that was used to build a prediction model) and a test set (that was used to test the model's predictive ability). The training set was composed of 160 randomized spectra (consisting of 80 normal and 80 colorectal cancer subjects), and the test set was composed of the remaining 46 spectra (consisting of 23 normal and 23 colorectal cancer subjects). The PLS approach was performed on the normalized spectral data using "Unscrambler" Version 9.7 (CAMO Software AS, Trondheim, Norway).

3 Results and Discussion

3.1 Nanoparticle Aggregation for Additional SERS Enhancement

As illustrated in Fig. 1(a), the membrane containing proteins was separated into two parts: half of the membrane was used as a reference and the other half was applied for the SERS detection. Using a part of the membrane itself as a reference can obviously improve the accuracy and efficiency of protein separation. However, the protein content for SERS detection is reduced to half of its original, at a concentration of approximately 0.42 to 0.6 g/L for albumin and 0.23 to 0.5 g/L for globulin. To achieve a high signal-to-noise ratio available for cancer detection, the SERS signal should be amplified at least twice as much as the original signal. Aggregating agents are routinely added to nanoparticles to increase the magnitude of the SERS enhancement.[28,29] However, in some cases, bands from aggregating agents such as KCl and KNO_3 may interfere with the SERS bands of the analyte.[28,29] By contrast, the use of acetic acid as the aggregating agent can avoid this problem. The flat, near-zero background signal of the mixture of acetic acid and the blank membrane has been demonstrated in our previous work.[22]

The aggregation protocol in Fig. 1(b) depicts the main factors that contribute to the additional SERS enhancement. The first factor is the increase in protein-nanoparticle interaction. The isoelectric point (pI) of the serum protein is in the range of 5 to 7. At a low pH condition (lower than pI of serum protein), serum proteins will carry net positive charges. Since the hydroxylamine-reduced Ag NPs have negative charges,[30] serum protein will absorb on Ag surfaces by electrostatic interaction. Furthermore, multiple interaction sites of an individual protein may bridge two or more Ag NPs and induce a subsequent SERS enhancement.[31] The second factor is the increase in nanoparticle-nanoparticle interaction. The existence of H^+ can neutralize parts of negative-charged Ag NPs and consequently reduce the repulsion between Ag NPs, which may promote the aggregation of Ag NPs to some extent.

To characterize the average particle size or aggregation effect, the position of the maximum absorption detected by UV-vis spectroscopy was recorded.[24] Figure 1(c) shows the UV-vis absorption spectra of the Ag NPs, albumin-Ag NPs mixture, and globulin-Ag NPs mixture. The maximum absorption of

Ag NPs is 417 nm, shifting to 426 nm for the albumin-Ag NPs mixture, and 428 nm for the globulin-Ag NPs mixture. The band shift demonstrates that aggregation of serum protein-Ag NPs occurs to some extent.

Figure 2(a) displays the average SERS spectra of albumin from the normal group ($n = 103$) and the colorectal cancer group ($n = 103$) together with standard deviations (SDs). Figure 2(b) shows the similar type of data for globulin. Prominent SERS bands of albumin and globulin are observed as follows: 522, 620, 644, 760, 830, 851, 880/883, 936, 1006, 1029, 1046, 1205, 1263, 1313/1324, 1447, 1551, and 1683 cm^{-1}.[22,32–39] Tentative assignments of these peaks are summarized in Table 2. As shown in Figs. 2(a) and 2(b), the distinct amide regions, tryptophan and tyrosine, as well as the SS region indicate the existence of secondary structures of serum proteins in acidic pH. Additionally, each small SD shows that the intersubject variations are relatively subtle. This may be because of the removal of undesired materials in the serum sample.
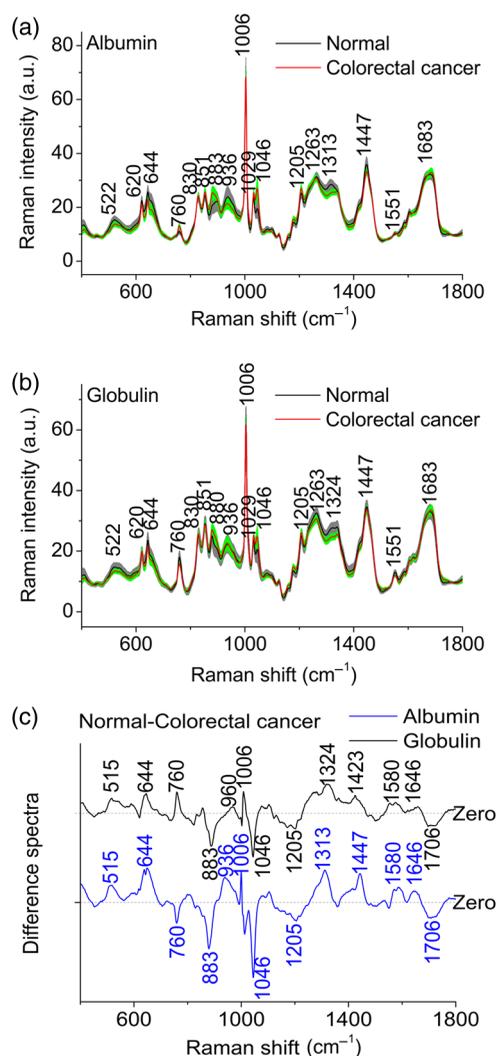

Fig. 2 Average SERS spectra of (a) albumin and (b) globulin from the normal group ($n = 103$) and the colorectal cancer group ($n = 103$), respectively. The shaded areas represent the standard deviations of the means. (c) Albumin and globulin difference spectra obtained by subtracting the colorectal cancer group from the normal group.

Table 2 Peak assignments for serum protein SERS spectra.[22,32–42,43,44]

Peak position (cm^{-1})	Major assignments
515/522	S—S stretching
557	Tryptophan
620	Phenylalanine
644	Tyrosine
690	C—S stretching
716	Methionine
760	Tryptophan
830	Tyrosine
851	Tyrosine
880/883	Tryptophan
936	C—C stretching
960	C—C stretching
1006	Phenylalanine
1029	Phenylalanine
1046	Tryptophan
1205	Phenylalanine, tyrosine
1263	Amide III
1313/1324	Amide III
1423/1447	CH$_2$ bending or scissoring
1551	Tryptophan
1580	Tryptophan, phenylalanine
1646/1683/1706	Amide I

3.2 SERS Spectra of Serum Proteins

The difference spectra shown in Fig. 2(c) are calculated by subtracting the average SERS spectra of the colorectal cancer group from that of the normal group. The blue line is the albumin difference spectrum and the black line is the globulin difference spectrum. Distinct spectral changes (e.g., Raman peak intensities, Raman peak positions, and spectral bandwidths broadening or narrowing) can be observed in both difference spectra around 515, 644, 760, 883, 936/960, 1006, 1046, 1205, 1313/1324, 1423/1447, 1580, 1646, and 1706 cm^{-1}.[32,36,38–42] The peaks assigned to phenylalanine (1006 and 1580 cm^{-1}) and tyrosine (644 cm^{-1}) show lower intensities in the colorectal cancer group, while the peak at 1205 cm^{-1}, which is related to both phenylalanine and tyrosine, is higher in the colorectal cancer group. Because 644, 1006, 1205, and 1580 cm^{-1} belong to the C—C twisting mode of tyrosine, the symmetric ring breathing mode of phenylalanine, the ring breathing of phenylalanine or tyrosine, and the C=C stretching of phenylalanine,

respectively,[32,41] it is possible that even though (1006, 1205, and 1580 cm^{-1}) or (1205 and 644 cm^{-1}) are assigned to the same kind of biomolecules, their molecular vibration modes are different. The band around 1580 cm^{-1} also belongs to tryptophan, but other tryptophan bands (883 and 1046 cm^{-1}) in the difference spectra show higher intensities in the colorectal cancer group. It is supposed that the 1580 cm^{-1} band in the spectra of serum proteins is dominated by the vibration of phenylalanine. Actually, most of the bands in the two difference spectra have similar profiles except for the peak at 760 cm^{-1}. The intensity of the 760 cm^{-1} band is negative in the albumin difference spectrum, while it is positive in the globulin difference spectrum. This means that the content of tryptophan existing in the albumin increases, but the content in the globulin decreases with cancer development. Some earlier findings have reported elevated levels of plasma free tryptophan in animals[45] and patients[46] with cancer, particularly those with anorexia.[47,48] However, in 2002, Huang et al. found a statistically significant 16% lowering of serum tryptophan among colorectal cancer patients compared with "no cancer" controls,[49] and this was compatible with some other results reported for patients with a variety of cancer types.[50,51] The differences between these results deserve further investigation due to the importance of tryptophan catabolism in the immunobiology of cancer.

The 1646 and 1706 cm^{-1} peaks are attributed to the α-helix and β-sheet structures of amide I, respectively.[52] The 1646 cm^{-1} peak is higher, but 1706 cm^{-1} peak is lower in the normal group than in the colorectal cancer group, suggesting that the colorectal cancer group has fewer α-helix but more β-sheet structural elements than the normal group. The structure of albumin is very flexible and it readily changes shape with variations in environmental conditions and with the binding of ligands;[15] as a result, changes in the albumin structure could indicate changes in the types of biological substances the albumin is carrying. As for globulin, the increase in β-sheet structural elements could very much indicate increased immunoglobulins, whose folds are mostly composed of β-sheet secondary structures and one disulfide bond. The conformational change of α-helix to β-sheet structure has also been reported for other diseases, such as melanoma, nasopharyngeal, and gastric cancer.[21,22,53]

3.3 PCA-LDA Analysis of SERS Spectra

To test the capability of serum proteins for distinguishing the colorectal cancer group from the normal group, a multivariate statistical method based on the combination of PCA and LDA was performed on the normalized SERS spectra. Figure 3 shows the trends of the classification accuracy (black line) and cumulative variance explained (blue line) with the varying number of PCs. Interestingly, the classification accuracy and cumulative variance explained are greatly improved when the number of PCs increases from 1 to 3. When the first three PCs are projected to LDA, a classification accuracy of ~97% can be achieved in both cases. The successive PCs describe the spectral features that contribute progressively smaller variances and accuracy increments. In some points, a dip [such as PCs = 4 in Fig. 3(b)] in accuracy may even appear. This may be because some PCs represent the background signal or common characterizations between the normal group and the cancer group. Hence, the combination of these PCs for PCA-LDA analysis does not guarantee an improvement in accuracy. The data points marked by the red arrows indicate the highest classification accuracies of 100% achieved at PCs = 6 for
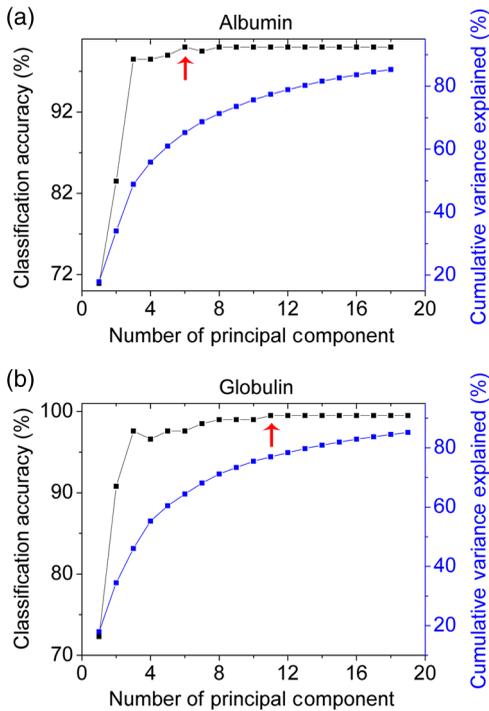
Fig. 3 Classification accuracy (black line) and cumulative variance explained (blue line) as a function of the number of principal components (PCs) for principal component analysis combined with linear discriminant analysis. The results are calculated from the dataset of (a) albumin and (b) globulin SERS spectra. The data points labeled by the red arrows correspond to the highest classification accuracies of 100% achieved at PCs = 6 for albumin and of 99.5% achieved at PCs = 11 for globulin.

Fig. 4 (a) Loadings of the first six PCs calculated from the albumin SERS spectral dataset. (b) Loadings of the first 11 PCs calculated from the globulin SERS spectral dataset.

albumin and of 99.5% achieved at PCs = 11 for globulin. The loadings of the first six PCs obtained from the albumin spectral dataset and the loadings of the first 11 PCs calculated from the globulin spectral dataset have been plotted in Fig. 4. Even though the latter PCs only account for small variances, they may also contain some information helpful for recognition and their removal may introduce a loss of discriminative information. For example, PC6 (~4.3%) and PC11 (~1.5%) only contain a small part of the spectral variances in albumin and globulin SERS spectra, respectively, but they promote accuracy improvement. The loadings of PC6 [Fig. 4(a)] show intense positive features at 832 and 1210 cm^{-1}, indicating different contents of tyrosine and phenylalanine in the albumin obtained from the normal group and the colorectal cancer group. The loadings of PC11 [Fig. 4(b)] are dominated by the vibrational feature of methionine with a positive band around 716 cm^{-1}. Both PC6 and PC11 provide a good complement to the variances explained by other PCs.

The classification results of the PCA-LDA diagnostic models are shown in Fig. 5. As seen from Fig. 5(a), the discriminant scores of the normal group and the colorectal cancer group do not overlap. Hence, the PCA-LDA diagnostic model based on albumin SERS spectroscopy provides a diagnostic sensitivity and specificity of 100% for discriminating the normal group from the colorectal cancer group. For globulin, a clear separation between the normal group and the colorectal cancer group can be observed with only one case of misclassification: one spectrum from the normal group is incorrectly classified into the colorectal cancer group. Therefore, the PCA-LDA
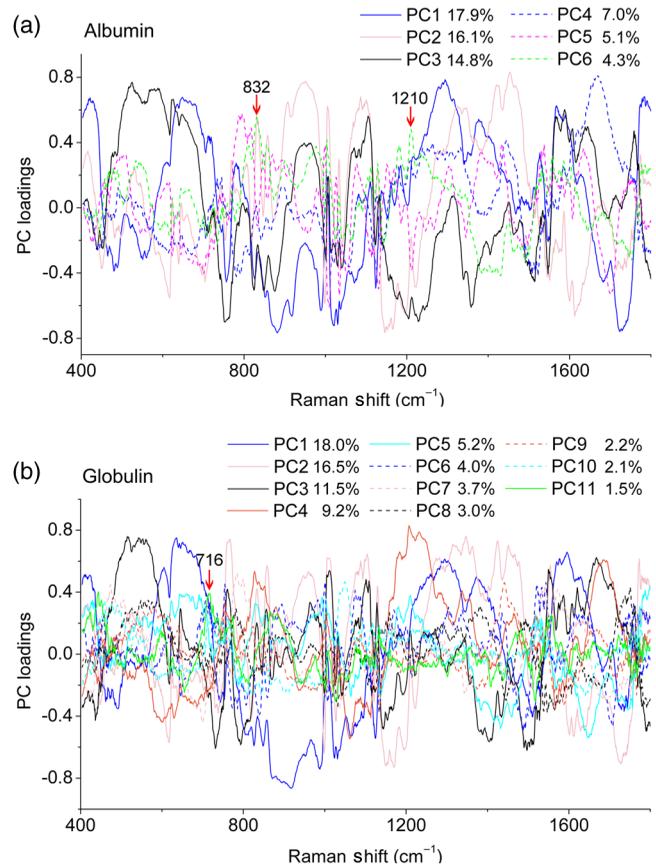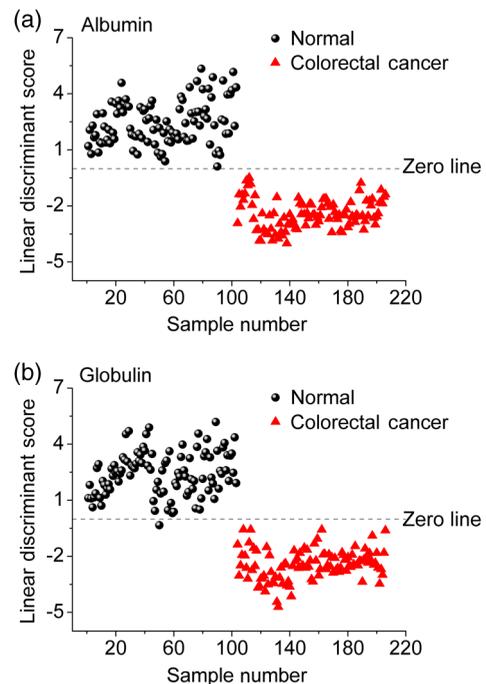

Fig. 5 Scatter plots of the linear discriminant scores belonging to the normal and colorectal cancer group, calculated from the datasets of (a) albumin and (b) globulin SERS spectroscopy.

diagnostic model based on globulin SERS spectroscopy provides a sensitivity of 100% and a specificity of 99% in the diagnosis of the colorectal cancer.

In both of the PCA plots, a number of samples remain close to the zero line, probably because some cancer patients showed a normal level of cancer biomarkers. For instance, it has been reported that only 4% of patients with stage T1 had an elevated carcinoembyonic antigen (>5 ng/mL), whereas 25%, 44%, and 65% of patients with stages T2, T3, and T4, respectively, had abnormal levels.[54] The next step, therefore, will be the study of the characteristic Raman peaks of these cancer biomarkers which play a key role in discriminating colorectal cancer patients from healthy controls.

We also examined the influence of the factors of age and gender on the diagnostic accuracy. No significant differences in the SERS spectra of the serum proteins purified from different age groups or different gender groups of the same pathology type can be observed. Interestingly, the SERS data seem to have some potential correlations with cancer stages. In the next step, we will concentrate on more detailed and prospective studies, including assessing samples from different pathology conditions to test whether this diagnostic platform has the ability to discriminate subpathology classes within colorectal cancer pathology and other cancer types (e.g., gastric cancer and esophagus cancer).

3.4 PLS Models for Predicting Colorectal Cancer

Given the excellent performance of PCA-LDA in discriminating the colorectal cancer from the normal subjects, the PLS approach was employed to demonstrate the predictive power of the proposed approach. Figures 6(a) and 6(c) provide the results of the leave-one-out cross-validation analysis for the albumin and globulin PLS models, respectively. In particular, the albumin PLS model is comprised of the first three LVs and the globulin PLS model consists of the first four LVs. The number of LVs used in the PLS model is automatically determined by the decision rule of the Unscrambler package together with a cross-validation analysis. Both figures plot the predicted values on the y-axis and the reference values on the x-axis. Relevant figures have also been summarized on the corresponding plots. R values for albumin and globulin PLS models are 0.944 and 0.922, respectively, exhibiting good fitting and predictive capacities. The RMSE values for both models (0.316 for albumin and 0.366 for globulin) are not satisfactory, which may be due to the small size of the training set.

In order to put the predictive capability of our models in perspective, the reliability of the predictions on "unknown" samples was validated in Figs. 6(b) and 6(d). The zero line is the boundary of the normal group and the colorectal cancer group, which is determined according to class membership (normal group: 1; colorectal cancer group: −1). As seen from the plots, the albumin PLS model provides a diagnostic accuracy of 93.5% (43/46) [sensitivity of 95.6% (22/23) and specificity of 91.3% (21/23)], and the globulin PLS model yields a diagnostic accuracy of 93.5% (43/46) [sensitivity of 91.3% (21/23) and specificity of 95.6% (22/23)] for colorectal cancer detection. However, some points as well as their relative SDs are near or even over the zero line, suggesting the potential possibilities might be misidentified.

4 Conclusions

We have studied colorectal cancer using serum protein SERS spectroscopy. Acetic acid, as a new aggregating agent, increases the magnitude of the SERS enhancement without contaminating the serum protein SERS signal. The difference spectra of serum
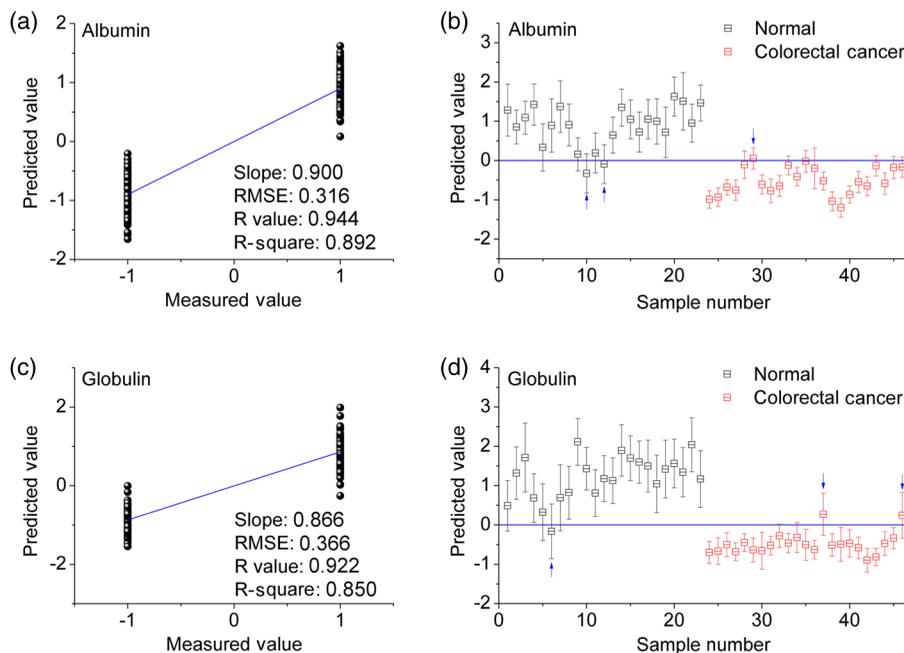

Fig. 6 Partial least squares (PLS) prediction models for correct classification of normal and colorectal cancer. Training sets were calculated from (a) albumin (160 spectra consisting of 80 normal subjects and 80 colorectal cancer subjects) and (c) globulin SERS spectra (160 spectra consisting of 80 normal subjects and 80 colorectal cancer subjects) and validated by leave-one-out cross-validation method. PLS prediction results based on (b) albumin and (d) globulin PLS prediction models. The blue arrows indicate the misjudged samples.

proteins and the corresponding PC loading plots, calculated from the spectral dataset of the colorectal cancer group and the normal group, demonstrate that the secondary structures of serum proteins and the contents of amino acids (e.g., tryptophan) change during cancer progression. PCA-LDA analysis and the PLS approach show that SERS spectra of albumin and globulin can provide a rapid and sensitive "Yes/No" assessment to identify the colorectal cancer.

Acknowledgments

References

1. J. Piva et al., "Overview of the use of theory to understand infrared and Raman spectra and images of biomolecules: colorectal cancer as an example," *Theor. Chem. Acc.* **130**(4–6), 1261–1273 (2011).
2. A. Jemal et al., "Global cancer statistics," *CA Cancer J. Clin.* **61**(2), 69–90 (2011).
3. R. A. Smith et al., "American Cancer Society Guidelines for the early detection of cancer: update of early detection guidelines for prostate, colorectal, and endometrial cancers: ALSO: update 2001—testing for early lung cancer detection," *CA Cancer J. Clin.* **51**(1), 38–75 (2001).
4. S. J. Winawer, "Colorectal cancer screening," *Best. Pract. Res. Clin. Gastroenterol.* **21**(6), 1031–1048 (2007).
5. P. Zirak et al., "Transcranial diffuse optical monitoring of microvascular cerebral hemodynamics after thrombolysis in ischemic stroke," *J. Biomed. Opt.* **19**(1), 018002 (2014).
6. F. El-Ghussein et al., "Hybrid photomultiplier tube and photodiode parallel detection array for wideband optical spectroscopy of the breast guided by magnetic resonance imaging," *J. Biomed. Opt.* **19**(1), 011010 (2014).
7. M. S. Schmidt et al., "Trends in melanosome microcavitation thresholds for nanosecond pulse exposures in the near infrared," *J. Biomed. Opt.* **19**(3), 035003 (2014).
8. W. R. Premasiri, J. C. Lee, and L. D. Ziegler, "Surface-enhanced Raman scattering of whole human blood, blood plasma, and red blood cells: cellular processes and bioanalytical sensing," *J. Phys. Chem. B* **116**(31), 9376–9386 (2012).
9. L. Mavarani et al., "Spectral histopathology of colon cancer tissue sections by Raman imaging with 532 nm excitation provides label free annotation of lymphocytes, erythrocytes and proliferating nuclei of cancer cells," *Analyst* **138**(14), 4035–4039 (2013).
10. V. Ranc et al., "Discrimination of circulating tumor cells of breast cancer and colorectal cancer from normal human mononuclear cells using Raman spectroscopy," *Analyst* **138**(20), 5983–5988 (2013).
11. X. Li, T. Yang, and S. Li, "Discrimination of serum Raman spectroscopy between normal and colorectal cancer using selected parameters and regression-discriminant analysis," *Appl. Opt.* **51**(21), 5038–5043 (2012).
12. C. Krafft et al., "Raman and FTIR microscopic imaging of colon tissue: a comparative study," *J. Biophotonics* **1**(2), 154–169 (2008).
13. G. B. Mider, E. L. Alling, and J. J. Morton, "The effect of neoplastic and allied diseases on the concentrations of the plasma proteins," *Cancer* **3**(1), 56–65 (1950).
14. D. Gupta and C. G. Lis, "Pretreatment serum albumin as a predictor of cancer survival: a systematic review of the epidemiological literature," *Nutr. J.* **9**(69) (2010).
15. J. Nicholson, M. Wolmarans, and G. Park, "The role of albumin in critical illness," *Br. J. Anaesth.* **85**(4), 599–610 (2000).
16. V. Fritz and L. Fajas, "Metabolism and proliferation share common regulatory pathways in cancer cells," *Oncogene* **29**(31), 4369–4377 (2010).
17. E. T. Fung et al., "Classification of cancer types by measuring variants of host response proteins using SELDI serum assays," *Int. J. Cancer* **115**(5), 783–789 (2005).
18. Y. Chen et al., "Label-free serum ribonucleic acid analysis for colorectal cancer detection by surface-enhanced Raman spectroscopy and multivariate analysis," *J. Biomed. Opt.* **17**(6), 067003 (2012).
19. D. Lin et al., "Colorectal cancer detection by gold nanoparticle based surface-enhanced Raman spectroscopy of blood serum and statistical analysis," *Opt. Express* **19**(14), 13565–13577 (2011).
20. A. Bonifacio et al., "Surface-enhanced Raman spectroscopy of blood plasma and serum using Ag and Au nanoparticles: a systematic study," *Anal. Bioanal. Chem.* **406**(9–10), 2355–2365 (2014).
21. J. Lin et al., "Surface-enhanced Raman scattering spectroscopy for potential noninvasive nasopharyngeal cancer detection," *J. Raman Spectrosc.* **43**(4), 497–502 (2012).
22. J. Lin et al., "A novel blood plasma analysis technique combining membrane electrophoresis with silver nanoparticle based SERS spectroscopy for potential applications in non-invasive cancer detection," *Nanomedicine* **7**(5), 655–663 (2011).
23. J. Wang et al., "Serum albumin and globulin analysis for hepatocellular carcinoma detection avoiding false-negative results from alpha-fetoprotein test negative subjects," *Appl. Phys. Lett.* **103**(20), 204106 (2013).
24. N. Leopold and B. Lendl, "A new method for fast preparation of highly surface-enhanced Raman scattering (SERS) active silver colloids at room temperature by reduction of silver nitrate with hydroxylamine hydrochloride," *J. Phys. Chem. B* **107**(24), 5723–5727 (2003).
25. J. Zhao et al., "Automated autofluorescence background subtraction algorithm for biomedical Raman spectroscopy," *Appl. Spectrosc.* **61**(11), 1225–1232 (2007).
26. S. Wold, K. Esbensen, and P. Geladi, "Principal component analysis," *Chemom. Intell. Lab. Syst.* **2**(1–3), 37–52 (1987).
27. J. Ghasemi, S. Ahmadi, and K. Torkestani, "Simultaneous determination of copper, nickel, cobalt and zinc using zincon as a metallochromic indicator with partial least squares," *Anal. Chim. Acta* **487**(2), 181–188 (2003).
28. T. W. Koo et al., "Specific chemical effects on surface-enhanced Raman spectroscopy for ultra-sensitive detection of biological molecules," *Appl. Spectrosc.* **58**(12), 1401–1407 (2004).
29. N. R. Yaffe and E. W. Blanch, "Effects and anomalies that can occur in SERS spectra of biological molecules when using a wide range of aggregating agents for hydroxylamine-reduced and citrate-reduced silver colloids," *Vib. Spectrosc.* **48**(2), 196–201 (2008).
30. X. Dong, H. Gu, and F. Liu, "Study of the surface-enhanced Raman spectroscopy of residual impurities in hydroxylamine-reduced silver colloid and the effects of anions on the colloid activity," *Spectrochim. Acta Part A* **88**, 97–101 (2012).
31. X. X. Han et al., "Label-free highly sensitive detection of proteins in aqueous solutions using surface-enhanced Raman scattering," *Anal. Chem.* **81**(9), 3329–3333 (2009).
32. K. W. Poon et al., "Quantitative reagent-free detection of fibrinogen levels in human blood plasma using Raman spectroscopy," *Analyst* **137**(8), 1807–1814 (2012).
33. N. C. Dingari et al., "Raman spectroscopy provides a powerful diagnostic tool for accurate determination of albumin glycation," *PLoS One* **7**(2), e32406 (2012).
34. G. V. P. Kumar et al., "Surface-enhanced Raman scattering studies of human transcriptional coactivator p300," *J. Phys. Chem. B* **110**(33), 16787–16792 (2006).
35. S. Stewart and P. Fredericks, "Surface-enhanced Raman spectroscopy of peptides and proteins adsorbed on an electrochemically prepared silver surface," *Spectrochim. Acta Part A* **55**(7–8), 1615–1640 (1999).
36. C. S. Teixeira et al., "Thyroid tissue analysis through Raman spectroscopy," *Analyst* **134**(11), 2361–2370 (2009).
37. T. Kitagawa, T. Azuma, and K. Hamaguchi, "The Raman spectra of Bence-Jones proteins. Disulfide stretching frequencies and dependence of Raman intensity of tryptophan residues on their environments," *Biopolymers* **18**(2), 451–465 (1979).
38. M. Moreno et al., "Raman spectroscopy study of breast disease," *Theor. Chem. Acc.* **125**(3), 329–334 (2010).

39. X. Yang et al., "Highly sensitive detection of proteins and bacteria in aqueous solution using surface-enhanced Raman scattering and optical fibers," *Anal. Chem.* **83**(15), 5888–5894 (2011).

40. N. Howell and E. Li-Chan, "Elucidation of interactions of lysozyme with whey proteins by Raman spectroscopy," *Int. J. Food Sci. Technol.* **31**(5), 439–451 (1996).

41. X. X. Han et al., "Analytical technique for label-free multi-protein detection based on Western blot and surface-enhanced Raman scattering," *Anal. Chem.* **80**(8), 2799–2804 (2008).

42. E. S. Grabbe and R. P. Buck, "Surface-enhanced Raman spectroscopic investigation of human immunoglobulin G adsorbed on a silver electrode," *J. Am. Chem. Soc.* **111**(22), 8362–8366 (1989).

43. A. Sengupta, M. L. Laucks, and E. J. Davis, "Surface-enhanced Raman spectroscopy of bacteria and pollen," *Appl. Spectrosc.* **59**(8), 1016–1023 (2005).

44. N. Masuda et al., "Raman spectroscopic markers for the Xenopus Laevis oocyte expression system," *Curr. Phys. Chem.* **3**(2), 151–154 (2013).

45. R. Krause et al., "Brain tryptophan and the neoplastic anorexia-cachexia syndrome," *Cancer* **44**(3), 1003–1008 (1979).

46. A. Cascino et al., "Increased plasma free tryptophan levels in human cancer: a tumor related effect?," *Anticancer Res.* **11**(3), 1313–1316 (1991).

47. W. T. Chance, M. F. von Meyenfeldt, and J. E. Fischer, "Changes in brain amines associated with cancer anorexia," *Neurosci. Biobehav. Rev.* **7**(4), 471–479 (1983).

48. F. R. Fanelli et al., "Plasma tryptophan and anorexia in human cancer," *Eur. J. Cancer Clin. Oncol.* **22**(1), 89–95 (1986).

49. A. Huang et al., "Tryptophan and quality of life in colorectal cancer," *Adv. Exp. Med. Biol.* **527**, 353–358 (2003).

50. H. Denz et al., "Weight loss in patients with hematological neoplasias is associated with immune system stimulation," *Clin. Investig.* **71**(1), 37–41 (1993).

51. R. M. Giusti et al., "Differential patterns of serum biomarkers of immune activation in human T-cell lymphotropic virus type I-associated myelopathy/tropical spastic paraparesis, and adult T-cell leukemia/lymphoma," *Cancer Epidemiol. Biomarkers Prev.* **5**(9), 699–704 (1996).

52. Z. Chi et al., "UV resonance Raman-selective amide vibrational enhancement: quantitative methodology for determining protein secondary structure," *Biochemistry* **37**(9), 2854–2864 (1998).

53. M. Gniadecka et al., "Melanoma diagnosis by Raman spectroscopy and neural networks: structure alterations in proteins and lipids in intact cancer tissue," *J. Invest. Dermatol.* **122**(2), 443–449 (2004).

54. M. G. Fakih and A. Padmanabhan, "CEA monitoring in colorectal cancer. What you should know," *Oncology* **20**(6), 579–587 (2006).

Jing Wang is currently studying for her master's degree under the supervision of Rong Chen and Juqiang Lin. She is currently researching the application of SERS for biomedical study.

Duo Lin received his master's degree in physical electronics from Fujian Normal University. Currently, he is an assistant of the Fujian University of Traditional Chinese Medicine. His current research focuses on biomedical detection using SERS.

Juqiang Lin received the PhD degree of biomedical engineering from Huazhong University of Science and Technology in 2006. Then, he went to Utah State University as a visiting scholar (2009 to 2011). His research interests include Raman spectroscopy and biomedical photonics imaging.

Yun Yu received his master's degree from Fujian Normal University. Currently, he is an assistant of the Fujian University of Traditional Chinese Medicine. His current research focuses on SERS measurement and analysis of cell.

Zufang Huang obtained his PhD degree in medical spectroscopy and spectral analysis from Fujian Normal University in 2013. He is now mainly focusing on the diagnosis and quantitative analysis of human fluids by using Raman spectroscopy and SERS.

Yanping Chen obtained her PhD degree in tumor pathological diagnosis from Fujian Medical University. Her main interests are in applications of SERS-Immunoassay (SERSIA) to the pathologic detection and the laboratory diagnosis of tumor.

Jinyong Lin is currently researching the label-free detection of type 2 diabetes by Raman spectroscopy for his master's degree under the supervision of Rong Chen and Juqiang Lin.

Shangyuan Feng received his PhD degree from Fujian Normal University in 2011. He started a postdoctoral position at BC Cancer Agency, Canada, in 2013. Currently, he is an associate professor of the School of Optoelectronics and Information Engineering, Fujian Normal University, China. His research interest focuses on the application of SERS in biomedical diagnosis.

Buhong Li is a professor and vice dean with the School of Photonics and Electronic Engineering at Fujian Normal University. He earned his PhD degree in optical engineering from Zhejiang University in 2003. He was a visiting scientist at the University of Toronto from 2005 to 2007 and is currently a senior visiting fellow at Humboldt University of Berlin. His research interests include fluorescence spectroscopy and imaging, and the mechanism and dosimetry for photodynamic therapy.

Nenrong Liu received the MS degree in optics from the Fujian Normal University in 2006. She is currently working toward the PhD degree in optical engineering at Fujian Normal University. Her technical research is mainly about nonlinear spectroscopy and biomedical imaging for noninvasively early cancer diagnosis.

Rong Chen is a professor of the School of Optoelectronics and Information Engineering, Fujian Normal University, China. His research interest focuses on the application of Raman spectroscopy in biomedical detection.