Photodiagnosis and Photodynamic Therapy 39 (2022) 102949

Contents lists available at ScienceDirect

Photodiagnosis and Photodynamic Therapy

journal homepage: www.elsevier.com/locate/pdpdt

Surface-enhanced Raman spectroscopic analysis of centrifugally filtered blood serum samples of hepatitis C patients

Samra Shakeel [a], Haq Nawaz [a,*], Muhammad Irfan Majeed [a,*], Nosheen Rashid [b,*], Muhammad Rizwan Javed [c], Ayesha Tariq [a], Beenish Majeed [a], Aafia Sehar [a], Sania Murtaza [a], Nimra Sadaf [a], Gull Rimsha [a], Imran Amin [d]

[a] Department of Chemistry, University of Agriculture Faisalabad, Faisalabad 38000, Pakistan
[b] Department of Chemistry, University of Education, Faisalabad Campus, Faisalabad 38000, Pakistan
[c] Department of Bioinformatics and Biotechnology, Government College University Faisalabad (GCUF), Faisalabad 38000, Pakistan
[d] PCR Laboratory, PINUM Hospital, Faisalabad 38000, Pakistan

ARTICLE INFO

Keywords:
Surface-enhanced Raman spectroscopy
Centrifugal filters of 50 kda
Blood serum
Hepatitis C
Principal component analysis
Partial least square regression analysis

ABSTRACT

Background: Raman spectroscopy is able to analyze non-invasively, disease related to body fluids.
Objectives: For the qualitative and quantitative analysis of HCV serum samples surface-enhanced Raman spectroscopy (SERS) based method is developed.
Method: Surface-enhanced Raman spectroscopy (SERS) technique is employed for analysis of filtrate portions of blood serum samples of hepatitis C virus (HCV) infected patients and healthy ones by using 50 kDa centrifugal filter device. The filtrate portions of the serum obtained in this way contain proteins smaller than 50 kDa and removal of bigger size protein which allows to acquire SERS spectral features of smaller proteins more effectively which are probably associated with Hepatitis C infection. Moreover, SERS spectral features of the filtrates of different level of viral load including low, medium and high viral loads are compared with SERS spectral features of the filtrate portions of healthy/control serum samples. SERS spectral data sets of different samples are further analyzed by using multivariate data analysis techniques such as principal component analysis (PCA) and partial least square regression (PLSR). Some SERS spectral features are solely observed in the filtrate portions of the serum samples of hepatitis C and their intensities are increased as the level of viral load increases and might be used for HCV diagnosis.
Results: PCA was found helpful for differentiation of SERS spectral data sets of filtrate portions of the serum samples of hepatitis C and healthy persons. The PLSR model helped for the quantification of viral loads in the unknown serum samples with 99% accuracy.

1. Introduction

Hepatitis C is caused by RNA virus and is extremely transmittable disease, affecting millions of people worldwide [1–3]. HCV is related to Flaviviridae family of virus and produces unhealthy cells in the cytoplasm of hepatocytes. The poor response of immune system against the mutational spreading of HCV is one of the main reasons of viral evasion scheme. Mutation of HCV from the adaptive immune response is one of the major viral division strategies. The duplication rate is 10^{12} virions per day and can be transmitted from person to person by transplantation of physical relation, infected medical equipment through hemodialysis and circulatory drugs. There are two major types of HCV infection including longtime infection called chronic and a short span infection known as acute infection [4]. For the HCV diagnosis, different gold standard techniques are being used including serological assay [5], recognition of antibodies of HCV [6] and enzyme linked Immuno-sorbent assay (ELISA) [7]. Notably, the ELSIA technique can show false positive and false negative results [8]. On the other hand polymerase chain reaction (PCR) is also being used for the direct analysis of HCV and recognition of RNA sequence present in the blood serum samples [7]. PCR is a bio-molecular based method and can easily detect the HCV infection by using primers [7] but it requires many consumables hence costly, it is highly sensitive technique even any form of contamination of sample can give false results. Furthermore, PCR can

* Corresponding authors.
E-mail addresses: haqchemist@yahoo.com (H. Nawaz), irfan.majeed@uaf.edu.pk (M.I. Majeed), nosheenrasheed@yahoo.com (N. Rashid).

https://doi.org/10.1016/j.pdpdt.2022.102949
Received 11 March 2022; Received in revised form 12 May 2022; Accepted 1 June 2022
Available online 2 June 2022
1572-1000/© 2022 Elsevier B.V. All rights reserved.

only be used to identify the presence or absence of known pathogens or genes [9]. Therefore, there is a requirement of a technique which is reliable, cost effective and can do quick analysis for the identification of HCV infection [10]. Raman spectroscopy is an analytical technique which is being used for monitoring biochemical changes on the basis of spectral deviations. It has been explored for various applications including analysis of bio macromolecules containing lipids [11], proteins [12] and DNA/RNA [13]. Moreover, Raman spectral characterization of the hepatitis C related antigens is performed which provides basic information about disease [14].

Serum is present in the human body having all kind of the protein except fibrinogen [15]. Surface enhanced Raman spectroscopy of the serum samples of different diseases has been done .in previous different studies including breast cancer dengue fever [8], hepatitis B and C virus [16,17], malaria [18].

In order to analyze the blood serum samples more effectively, separation of proteins of different sizes from the serum can be done by using centrifugal filters of different KDa sizes with the help of centrifugal filters. In the current study, the technique of centrifugal filtration is used for the blood serum samples of Hepatitis C patients. Notably, our previous study is about application of the SERS technique for the hepatitis B (HBV) serum samples [19]. It is important to mention that both HBV and HCV are two different diseases. The HBV is a DNA based virus and HCV is an RNA based virus. In this way their respective SERS features are expected to be different from each other. This can lead to the development of the surface enhanced Raman spectroscopy-based methods for the identification of the viral infections in the blood serum samples of the patients.

This is the first time that we introduced the centrifugal filtration method to examine the qualitative and quantitative analysis of HCV serum samples which can potentially give more accurate results for the diagnosis of hepatitis C. Surface-enhanced Raman spectroscopy (SERS) employing silver nanoparticles as substrates, is used to analyze the filtrate portions of the serum samples of the HCV patients obtained by using centrifugal filters of 50 KDa. In this way, the filtrate portions of the serum samples contain proteins smaller than 50 KDa and the removal of bigger size proteins allows for the acquisition of SERS spectral features of smaller proteins more effectively which are probably associated with Hepatitis C infection. According to the best of our knowledge, there is no such work published yet.

2. Material and methods

2.1. Preparation of Ag NPs

Silver nanoparticles (Ag-NPs) were prepared by a chemical reduction process by using silver nitrate ($AgNO_3$) as a precursor and ($Na_3C_6H_5O_7$) trisodium citrate as a reducing agent and capping ligand. In this process 0.0085 g/300 ml of silver nitrate was added in deionized water and heated at 100 °C. As the boiling starts, 10 ml of 1% trisodium citrate was introduced to the solution. Keep the solution boiling for two to three hours on hot plate along with magnetic stirring to obtain gray colored silver nano-particles. The obtained Ag-NPs observed under the transmission and scanning electron microscopy. Both of these were performed by using a field emission electron microscope (JEOL, JSM-7500F) functioning in equally scanning and transmission modes at an accelerating voltage of 30 keV. In TEM, the nanoparticles are organized by slow evaporation method from aqueous suspension to from a drop of Ag-NPs on copper grid (400 mesh) coated with the carbon [20]. On the other hand, in SEM, dehydrated Ag-NPs are put on copper end lacking any additional conductive coating. ImageJ software was used for the analysis of the electron micrographs showing that the silver nanoparticles have oval geometry with an average size of 65 × 45 nm. These results are published elsewhere [17].

2.2. Sample collection and preparation

The clinically diagnosed HCV positive and healthy blood serum samples were acquired from the PINUM Hospital Faisalabad, Pakistan. The details of these samples are described in; Table 1. For study purpose, permission was taken from the Ethical Clearance Committee of University of Agriculture, Faisalabad, Pakistan. Notably, the for the purpose of understanding the serum samples were divided into three different levels of viral loads such as high, medium and low. There were no specific boundaries among these different groups of samples. The filtrates are obtained by filtering these serum samples using 50 KDa

Table 1.
Details of HCV positive serum samples categorized as low, medium, high viral loads, negative control and true healthy samples.

Series No.	Samples name	Viral load (IU)	Log of viral load	class
1	S2	1.91E+02	2.2810	Low
2	S8	2.60E+03	3.4149	Low
3	S21	6.64E+03	3.8221	Low
4	S24	1.26E+02	2.1003	Low
5	S27	3.03E+03	3.4814	Low
6	S14	7.40E+03	3.8692	Low
7	S52	4.57E+03	3.6599	Low
8	S77	4.52E+02	2.6551	Low
9	S93	2.35E+02	2.3710	Low
10	S34	6.22E+04	4.7937	Medium
11	S49	6.14E+04	4.7881	Medium
12	S75	5.89E+04	4.7701	Medium
13	S6	1.47E+05	5.1673	Medium
14	S32	7.55E+05	5.8779	Medium
15	S38	6.82E+05	5.8337	Medium
16	S43	2.63E+05	5.4199	Medium
17	S41	1.07E+05	5.0293	Medium
18	S49	3.45E+05	5.5378	Medium
19	S7	3.32E+05	5.5211	Medium
20	S54	6.83E+05	5.8344	Medium
21	S59	4.91E+06	6.6910	High
22	S9	3.64E+06	6.5611	High
23	S1	1.25E+06	6.0969	High
24	S51	2.04E+06	6.3096	High
25	S53	4.73E+06	6.6748	High
26	S48	1.52E+06	6.1818	High
27	S1	1.24E+06	6.0934	High
28	S29	1.08E+06	6.0334	High
29	S2	Negative Control	<1	Healthy
30	S3	Negative Control	<1	Healthy
31	S4	Negative Control	<1	Healthy
32	S7	Negative Control	<1	Healthy
33	S8	Negative Control	<1	Healthy
34	S9	Negative Control	<1	Healthy
35	S10	Negative Control	<1	Healthy
36	S11	Negative Control	<1	Healthy
37	S12	Negative Control	<1	Healthy
38	S13	Negative Control	<1	Healthy
39	S14	Negative Control	<1	Healthy
40	S15	Negative Control	<1	Healthy
41	S8	Negative Control	0	True Healthy
42	S16	Negative Control	0	True Healthy
43	S17	Negative Control	0	True Healthy

centrifugal filter and analysed by SERS technique. . The results of a viral load test are normally described in international units per milliliter (IU/ mL).

2.3. Centrifugal ultrafiltration of serum samples

This protocol of centrifugal filtration was adopted from the literature [21]. For the centrifugal filtration of the blood serum samples, 200 µl of each sample were put in the 50 KDa centrifuge tubes of Amicon Ultra-2 ml centrifugal devices (Merck, Millipore). The tubes were centrifuged at the speed of 6.5×10^3 RPM for 30 min. Same procedure is followed for all the hepatitis c and healthy serum samples. As a result of centrifugal filtration, 2 fractions were obtained one with molecular weight higher than 50 kDa (concentrate/ HMWF) called residue and other with molecular weight lower than 50 kDa (filtrate/ LMWF). These two fractions were collected in the separate vials.

2.4. Surface Enhance Raman spectroscopy (SERS) data acquisition

Surface Enhance Raman spectral data acquisition was done by using Raman Microscope Spectrometer (ATR8300BS) Optosky China with laser of 785 nm wavelength and a lens of 40X/0.6. For the spectral acquisition, 50 µl of each serum sample was mixed with 50 µl of Ag-NPs and left for 30 min for the incubation. From this incubation mixture, 50 µl were placed on the aluminum slide having a little groove and 15 SERS spectra were recorded. For each sample the laser power was set as100 mW and integration time was 10 s.

2.5. Data preprocessing

For the preprocessing of SERS spectral data, Matlab 7.8 already established chemometric protocols were used. Surface enhanced Raman spectral data sets of all the samples were put in a single matrix for the preprocessing steps including smoothing, substrate elimination, baseline correction and vector normalization. The smoothing of the SERS spectral data was done with the help of Savitzky-Golay smoothing technique (order 3, 17 points window). Notably, normalization algorithm is not applied for the quantitative analysis, Partial least square regression analysis (PLSR), in the current study The quantitative evidence maybe lost by applying normalization which can disturb the intensity and concentration relationship [39]. The baseline correction of every spectral data of SERS was modified by applying the rubber band alteration/ correction and the extra substances were also deducted from all the spectral data of every serum samples.

2.6. Data analysis

The peak assignments of the spectral features taken from literature were used for the understanding the reults as mentioned in the Table 2. For the SERS spectral data analysis, multivariate data analysis methods including principal component analysis (PCA) and Partial least square regression analysis (PLSR) are employed. PCA is a data analysis technique for unidentified samples for which previous information about the samples is not needed. PCA is an excellent technique for the identification of the variability present in the spectral data sets and it gives assessment about the sample data by decreasing dimensionality but keeping their variability. The maximum variability is explained by PC-1 and second maximum by PC-2 and so on [39].

For the PLSR analysis, the SERS spectral information is used as independent variables which are regressed against predetermined viral load values, known as dependent variables of the serum samples. The variance among the measured and unexplained predicted variables are residual variance. PLSR is used to analyze the serum samples quantitatively and predict the viral loads. In order to explain fitness of the model, root mean squared error of cross validation (RMSECV) and root mean squared error of prediction (RMSECV) were determined. Calibration was accomplished on 60% of randomly selected data sets and remaining 40% were utilized for the prediction/validation of the viral loads by using leave one sample out cross validation method [40].

Table 2.
RNA based Peak Assignment measured by SERS spectral data gained from literature. Citations are also given.

Wavenumber (cm^{-1})	Peak Assignments	Refs.
402	Cytosine stretching	[22]
436	Lipids	[23]
499	C-OH_3 torsion of methoxy	[24]
550	Lipids	[23]
573	cytosine	[22]
593	Phosphate of HA	[23]
596	Phosphatidylinositol	[23]
639	Amino acid	[25]
728	C—C stretching, proteins	[26]
743	T (ring breathing mode of thymine in DNA/RNA)	[27]
766	thymine ring breathing mode	[27]
817	C-C stretching (collagen assignment)	[26]
856	Amino acid	[28]
880	Tryptophan, sigma(ring)	[29]
888	Protein bands	[30]
945	stretching vibrations for the amino acids	[31]
1008	Phenylalanine (proteins)	[32]
1034	C—C stretching; proteins	[28]
1056	PO_2 stretching (DNA/RNA)	[33]
1100	n(C—C)-lipids, fatty acids	[23]
1102	Phenylalanine (proteins)	[33]
1133	Fatty acid	[23]
1209	Amide III-Due to C—N stretching and N—H Bending	[30]
1248	Amide III	[30]
1255	Lipids	[34]
1275	Amide III band in proteins	[25]
1332	-C stretch of Phenyl and C_3-C_3 stretch/in-plane bend of proteins	[35]
1333	Guanine	[22]
1374	RNA base	[25]
1379	lipid assignment	[36]
1411	Lipid bands	[36]
1450	CH_2 bending, CH functional groups in lipids, CH_2 bending (proteins)	[34]
1513	C-C stretching of cytosine	[34]
1533	Amide carbonyl group	[37]
1585	C=C olefinic stretch (protein assignment)	[26]
1590	Carbon particles	[28]
1626	Amide-I (a-helix)	[36]
1658	Amide-I (proteins)	[36]
1704	n(C_{55}O)OH (amino acids aspartic & glutamic acid)	[36]
1709	amino acids	[36]
1725	(C-O) stretching vibration in Guanine	[23]
1774	C-O and C-C stretching vibration (for Lipids and Fatty acids)	[38]

3. Results and discussion

Synthesized silver nanoparticles (Ag-NPs) were characterized by using transmission electron microscopy (TEM) which showed that Ag-NPs were found to be of oval shape with size ranging from 45 to 65 nm. Previous studies demonstrated that the nanoparticles of this size range were found to be excellent for SERS enhancement. In addition, use of centrifugal filers (50 KDa) can separate HCV disease associated proteins in the filtrate portion of the serum samples which can be further useful for the characterization of the biochemical changes as a result of development of disease. Different filtrate samples according to their predetermined viral loads, 43 filtrate samples, are classified into three categories such as high Viral load, medium viral load and low viral load samples and are compared against filtrates of healthy samples.

3.1. Mean SERS spectra

The peak assignments of these spectral features of HCV whole serum

samples, residue and filtrate were taken from the literature given in Table 2 where the citations are also mentioned against each peak assignment. Fig. 1 displays the mean spectra of centrifuged HCV serum samples which are separated into two components such as residual and filtrate. The solely differentiating features are represented by solid lines and intensities-based differences are labelled by dotted lines. In this mean spectral plot the different peaks labeled with solid lines include 573 cm^{-1} (Tryptophan/cytosine, guanine), 596 cm^{-1} (Phosphatidylinositol), 945 cm^{-1} (stretching vibrations for the amino acids), 1513 cm^{-1} (cytosine), 1626 cm^{-1} (amide I (alpha-helix)) and at 1725 cm^{-1} ((C—O) stretching vibration of guanine). On the other hand, the peaks represented by dotted lines having increase in intensities as the viral load increased, include 499 cm^{-1} (C—OH_3 torsion of methoxy), 639 cm^{-1} (n(C-S) gauche (amino acid methionine)), 743 cm^{-1} (T (ring breathing mode of thymine in DNA/RNA)), 817 cm^{-1} (C—C stretching (collagen assignment)), 888 cm^{-1} (Proteins), 1008 cm^{-1} (Phenylalanine), 1056 cm^{-1} (PO_2 stretching (DNA/RNA) chain, 1133 cm^{-1} (Fatty acid), 1209 cm^{-1} (amide III-Due to C—N stretching and N—H bending), 1333 cm^{-1} (guanine), 1411 cm^{-1} (vs. COO2), 1585 cm^{-1} (C=C olefinic stretch protein assignment), 1774 cm^{-1} (C—O and C—C stretching vibration of lipids and fatty acids). In the SERS spectrum of whole serum, these peaks are either absent or having less intensity and broad but in the spectra of residue or filtrate these are observed as new peaks or having increase in intensity indicating that due to centrifugal filtration the respective biomolecules are concentrated in the filtrate or residue as compared to whole serum. The peaks which are represented by dashed lines including 550 cm^{-1} (Cholesterol), 856 cm^{-1} (Amino acid side chain vibrations), 1100 cm^{-1} (n(C—C)-lipids, fatty acids), 1275 cm^{-1} (Amide III band in proteins), 1379 cm^{-1} (d(CH_3 symmetrical lipid assignment)), 1450 cm^{-1} (CH_2 bending) 1533 cm^{-1} (amide carbonyl group), 1658 cm^{-1} (amide-I (proteins), 1709 cm^{-1} (amino acids) decrease in intensity indicating that due to centrifugal filtration the respective biomolecules are eliminated from the filtrate or residue as compared to whole serum.. In the whole serum samples these peaks are intense but in residue and filtrate, they have decreased intensities.

After establishing the effectiveness of applying centrifugal filters (50 KDa) on the serum samples which is resulted in the form of different SERS spectra of whole serum, filtrate and residue (Fig. 1), this approach is used for different HCV positive serum samples having different levels of viral loads. These filtrate samples were analysed by using SERS technique. Fig. 2 shows mean SERS spectra of the filtrates of different HCV serum samples which are categorised into three groups including low, medium and high viral loads (Table 1). The significant spectral features which are labeled by solid lines including 402 cm^{-1} (Cytosine stretching), 596 cm^{-1} (phosphatidylinositol), 766 cm^{-1} (thymine ring breathing mode), 945 cm^{-1} (stretching vibrations of amino acids), 1513 cm^{-1} (C—C stretching of cytosine), 1725 cm^{-1} (C—O stretching vibration Guanine). The peaks which are represented by dotted lines have increasing intensities as the viral load of the filtrate samples increases including peaks at 499 cm^{-1} (C—OH_3 torsion of methoxy), 573 cm^{-1} (cytosine), 639 cm^{-1} (amino acid)), 817 cm^{-1} (C—C stretching of collagen assignment), 888 cm^{-1} (Protein bands), 1008 cm^{-1} (Phenylalanine), 1034 cm^{-1} (C—C stretching of protein), 1133 cm^{-1} (fatty acid),1209 cm^{-1} (Amide III-Due to C—N stretching and N—H Bending), 1374 cm^{-1} (RNA base), 1411 cm^{-1} (lipid bands), 1704 cm^{-1} (n($C_{55}O$) OH (amino acids aspartic & glutamic acid)) and 1774 cm^{-1}(Lipids and Fatty acids). In Fig. 2 the mean plot presents significant spectral features which are characterized by the dashed lines including 436 cm^{-1} (Cholesterol), 728 cm^{-1} (C—C stretching, proline (collagen assignment)), 856 cm^{-1} (Amino acid side chain vibrations), 1102 cm^{-1} (Phenylalanine (proteins)), 1255 cm^{-1} (Lipids), 1332 cm^{-1} (-HC stretch of Phenyl), 1450 cm^{-1} (CH_2 bending in lipids), 1590 cm^{-1} (Carbon particles), 1658 cm^{-1} (Amide-I (proteins). These characteristic peaks of lipids, proteins and DNA/RNA bases are the differentiating features in the Fig. 2 and can be considered as indications of the HCV virus infection.

The increasing intensities of RNA related SERS features in the filtrate samples can be due to the increase in the viral loads hence taken as indication of the HCV infection. The spectral feature present at 1209 cm^{-1} and 1135 cm^{-1} (amide-III Due to C—N stretching and N—H Bending) illustrate increased protein component in the filtrate sample which may be due to virus-related proteins. The peak observed at 1374 cm^{-1} (ring breathing modes in uracil) is directly linked with viral RNA because it is detected due to the uracil base that only exists in RNA. This is very significant spectral feature for identification of the Hepatitis virus because HCV is a RNA virus.

Other changes that are shown by dashed lines include features at 436, 728, 856, 1102, 1255, 1332, 1450, 1590 and 1658 cm^{-1} which decreased in their intensities as the viral load increases. At 436 cm^{-1} (Cholesterol, cholesterol ester), 728 cm^{-1} and 856 cm^{-1} (Amino acid

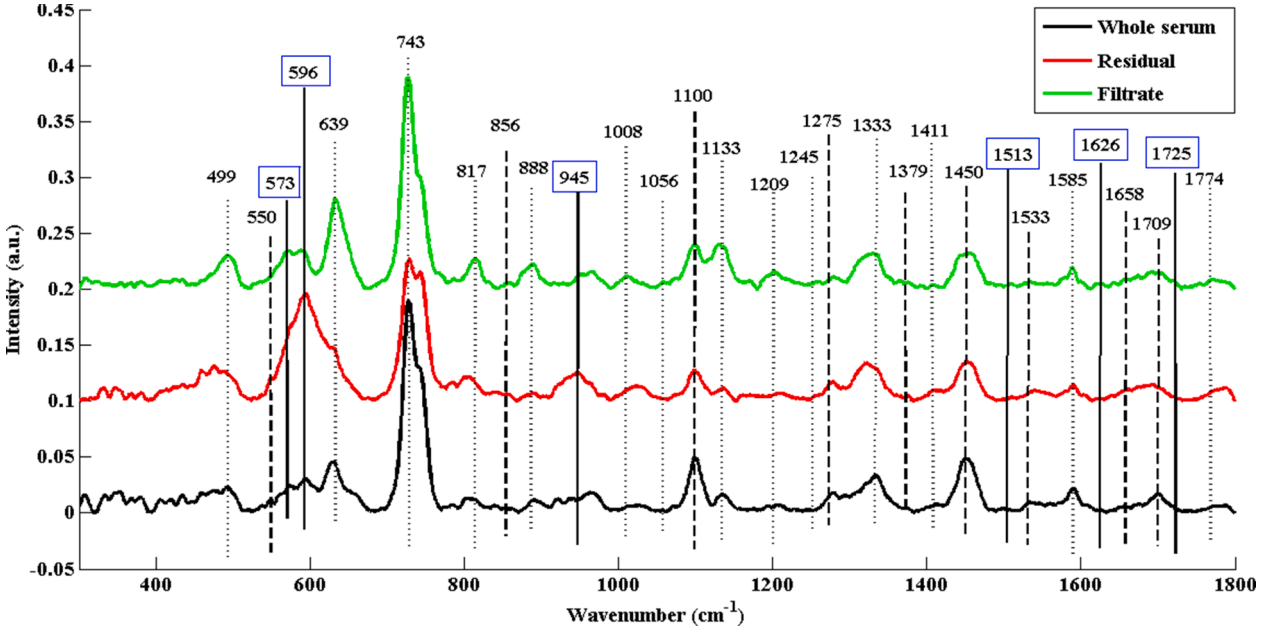


Fig. 1. Mean SERS spectra of whole serum, filtrate and residue portions obtained by use of centrifugal filters of 50 KDa.

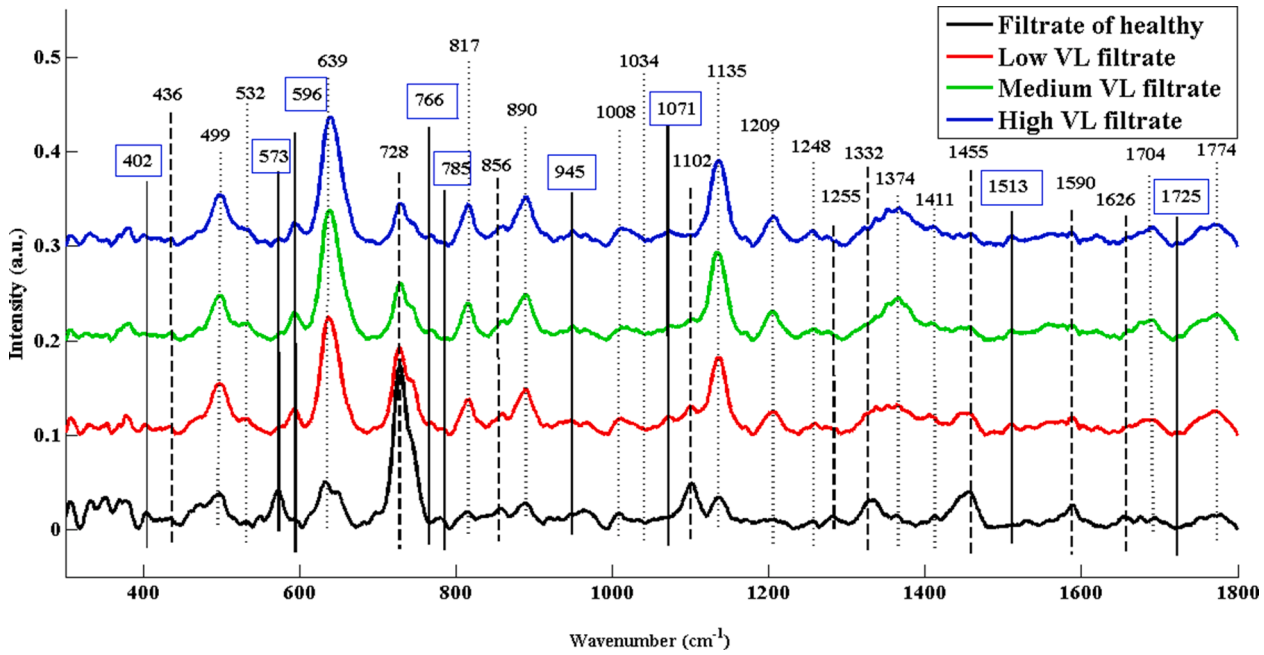


Fig. 2. Mean SERS spectra of filtrate portions of different HCV positive patients with increasing viral loads and healthy persons obtained by use of centrifugal filters of 50 KDa for filtering whole serum samples.

side chain vibrations of proline & hydroxyproline, as well as a (C—C) vibration of the collagen backbone Hydroxyproline (collagen type I)) and (C—C stretching, proline (collagen assignment)) are showing that at these points collagen contents are decreased as the viral load is increased. At 1102 cm^{-1} (phenylalanine (proteins)),1332 cm^{-1} (-C stretch of Phenyl (1) and C_3—C_3 stretch and C_5—O_5 stretch CH(alpha) in-plane bend), 1450 cm^{-1} (CH_2 bending of lipids), Furthermore, the SERS peak present at 1255 cm^{-1}, 1450 cm^{-1},1658 cm^{-1} (C=C stretching of lipids) show decreasing intensity as the viral load increases, indicating that lipid components in the serum are decreasing gradually due to increase of the severity of disease. It has been observed that the patients infected by HCV have considerably lower C—O and C—C stretching vibration (for Lipids and Fatty acids), cholesterol, triglyceride and lipoproteins [41].

3.2. Principal component analysis (PCA)

Principal component analysis (PCA) is applied to decrease the dimensionality of SERS spectral data in such a way that no variability is lost, in order to differentiate SERS spectral datasets of distinct groups of filtrates samples and whole HCV serum samples [7].

Fig. 3 illustrates the PCA scatter plot of SERS spectral data sets of different classes of HCV positive filtrates and whole serum samples as shown as different clusters differentiated from each other.

Three separate PCA scatter plots of SERS spectral data sets of filtrate portions of serum samples of “low VL” (9 samples), “medium VL” (11 samples) and “high VL” (8 samples) are shown in Fig. 4(a), Fig. 4(b) and Fig. 4(c) respectively. There is reasonable differentiation observed between SERS spectral data sets of filtrate samples of different viral loads.

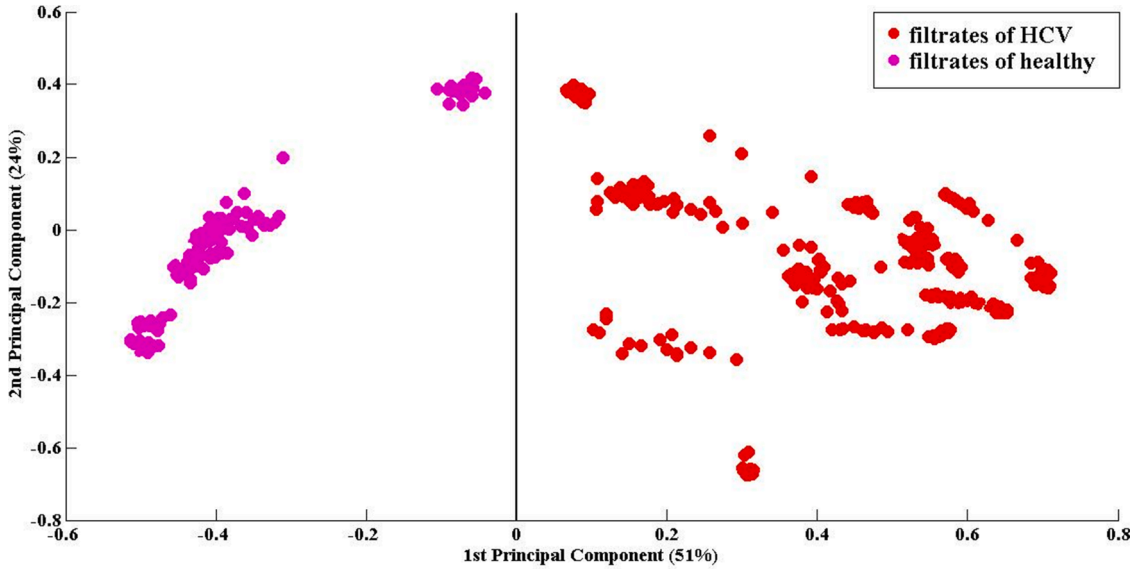


Fig. 3. PCA scattered plot of SERS spectral data sets of filtrate of healthy and filtrates of HCV samples collected from different patients having increasing viral loads obtained by use of centrifugal filters of 50 kDa.

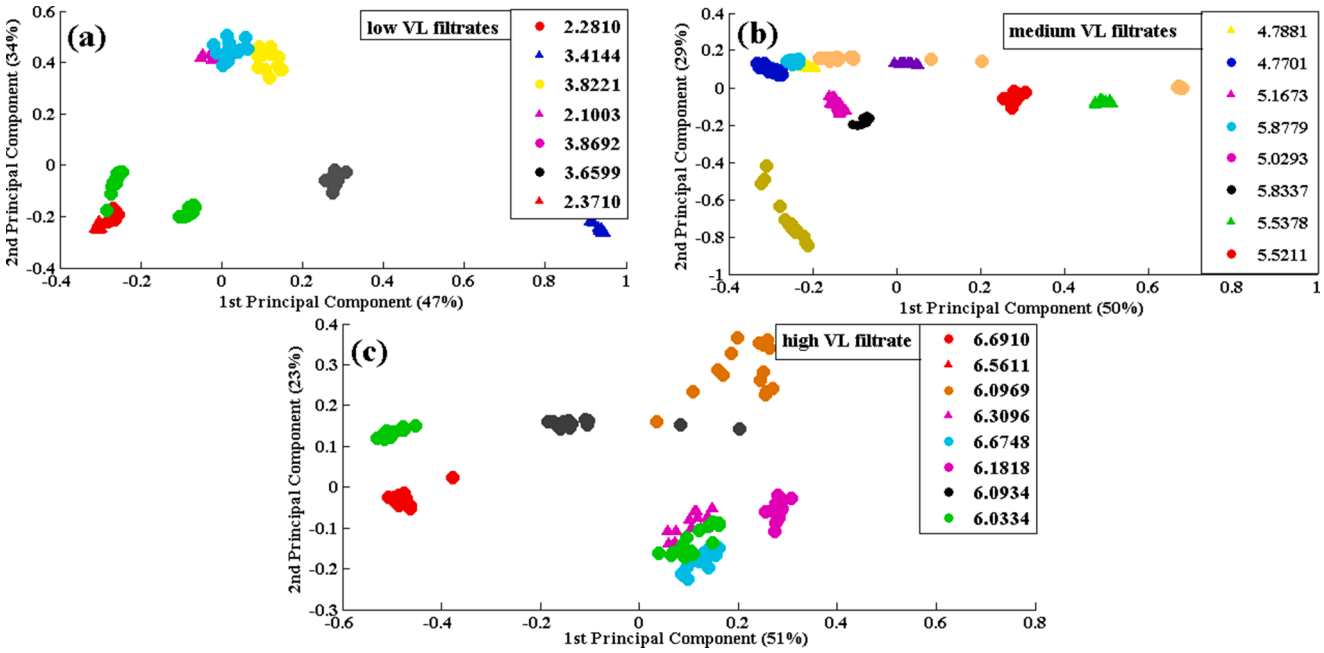


Fig. 4. PCA scatter plot of SERS spectral data sets of filtrate portions of (a) Low VL having 9 samples (b) Medium VL having 11 samples (c) High VL having 9 samples.

Moreover, for the identification of the loadings of PCA which are the basis of differentiation of different categories of filtrates having different viral load (low VL, medium VL and high VL), pairwise PCA analysis was performed as shown in Fig. 5 (a-i, b-i, c-i) as scatter plots and Fig. 5 (a-ii, b-ii, c-ii) as loadings.

Fig. 5a (i) shows PCA scatter plot of centrifuged filtrate of low viral load (9 samples) versus high viral load (8 samples) shown by red color and blue color dots, respectively. As shown in the plot that categories of low VL and high VL are distinguished on the basis of principal component analysis. For low vs. high viral load, PC-1 explains 47% of the variability in the two data sets and PC-2 explains 28%. The samples of filtrates of low viral load category are clustered on the negative side and the samples of filtrates of high viral load are clustered on the positive side which shows that samples of Low and high viral load are well separated due to difference of biochemical changes reflected in the form of SERS spectral features due to changes in viral load. Fig. 5a (ii) shows the loadings of PCA of categories of low vs high viral load samples. Remarkably, negative loadings are detected at 499 cm^{-1} (C—OH_3 torsion of methoxy), 639 cm^{-1} (amino acid), 817 cm^{-1} (C—C stretching collagen assignment), 888 cm^{-1} (protein band), 1008 cm^{-1} (phenylalanine), 1134 cm^{-1} (C—C stretching of protein), 1209 cm^{-1} (Amide III-Due to C—N stretching and N—H Bending), 1374 cm^{-1} (RNA base), 1411 cm^{-1} (lipid band), 1704 cm^{-1} (n(C_{55}O)OH (amino acids aspartic & glutamicacid) and 1774 cm^{-1} (lipid and fatty acid) which are associated

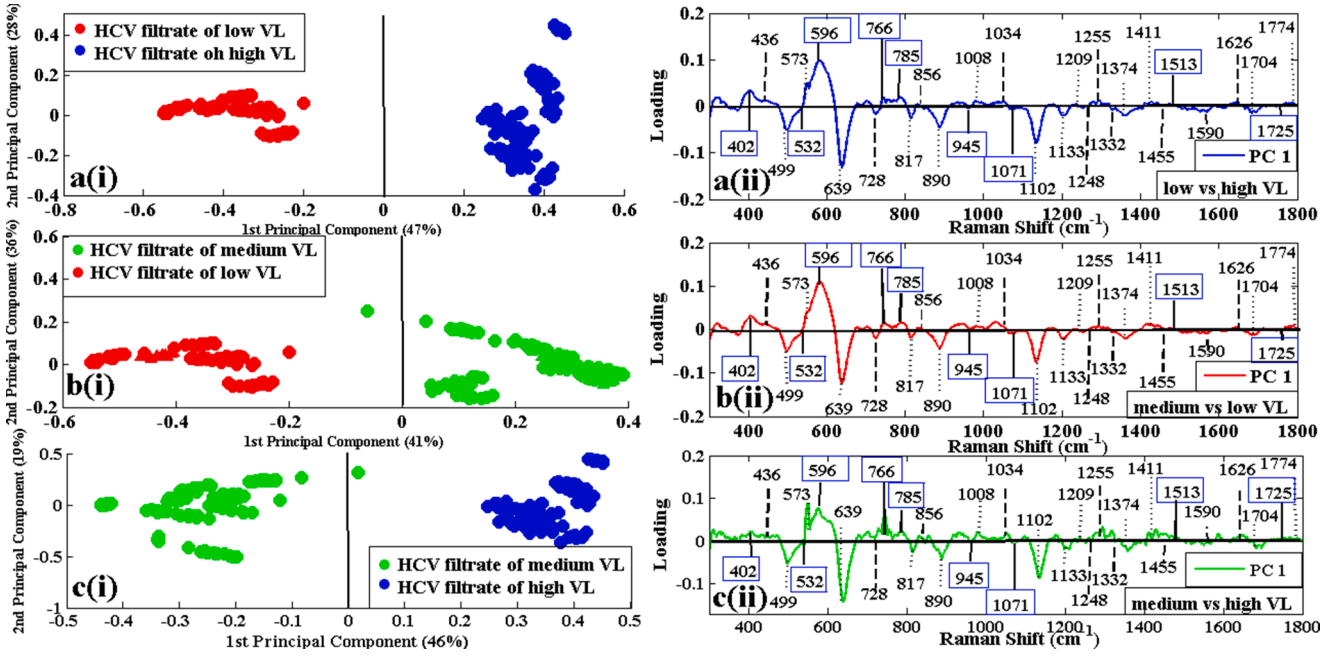


Fig. 5. Pairwise PCA scatter plots and loadings of SERS spectral data sets of filtrate portions of HCV positive serum samples (a-i and a-ii) low vs. high VL, (b-i and b-ii) low vs. medium VL and (c-i and c-ii) medium vs. high VL.

the SERS spectral data set of filtrates of high VL, clustered in the negative axis. These features indicate that as the viral load increases, intensities of these features increase. Other peaks that are observed on the positive side of PC-1 loadings contain stretching of PO_4^{-3} phosphate, Amide I and Amide-III β-sheet structure, C-C=C stretching in Adenine and C—O and C—C stretching lipids and fatty acids are all in accordance as the results described in mean spectra. The positive loadings of the filtrate samples including 436 cm^{-1} (Lipids), 728 cm^{-1} (C—C stretching, proteins), 856 cm^{-1} (Amino acid), 1102 cm^{-1} (Phenylalanine (proteins)), 1255 cm^{-1} (Lipids), 1332 cm^{-1} (-C stretch of Phenyl and C_3—C_3 stretch/in-plane bend of proteins), 1450 cm^{-1} (CH_2 bending, CH functional groups in lipids, CH_2 bending (proteins)), 1590 cm^{-1} (Carbon particles), and 1658 cm^{-1} (Amide-I (proteins))are associated with the SERS spectral data sets of low viral load class of samples and show decrease in intensities as the viral load increases.

Fig. 5b (i) shows PCA scatter plot of centrifuged filtrate samples of low viral load (9 samples) versus medium viral load (11 samples) shown by red color and green color dot, respectively. As shown in the plot these categories of low VL and medium VL are distinguished on the basis of principle component analysis. For low vs medium viral load, PC-1explains 41% of the variability in the two data sets and PC-2 explains 36%. The samples of the filtrates of low viral load category are the clustered on the negative side and the samples of filtrates of medium viral load are clustered on the positive side which shows that samples of low and medium viral load are well separated due to difference of biochemical changes reflected in the form of SERS spectral features due to changes in viral load. Fig. 5b (ii) shows the loadings of PCA of categories of low vs medium viral load samples. Remarkably, negative loadings are detected at 499 cm^{-1} (C—OH_3 torsion of methoxy), 639 cm^{-1} (amino acid), 817 cm^{-1} (C—C stretching collagen assignment), 888 cm^{-1} (protein band), 1008 cm^{-1} (phenylalanine), 1134 cm^{-1} (C—C stretching of protein), 1209 cm^{-1} (Amide III-Due to C—N stretching and N—H Bending), 1374 cm^{-1} (RNA base), 1411 cm^{-1} (lipid band), 1704 cm^{-1} (n(C_{55}O)OH (amino acids aspartic & glutamicacid) and 1774 cm^{-1} (lipid and fatty acid) which are associated the SERS spectral data set of filtrates of high VL, clustered in the negative axis. These features indicate that as the viral load increases, intensities of these features increase. Other peaks that are observed on the positive side of PC-1 loadings contain stretching of PO_4^{-3} phosphate, Amide I and Amide-III β-sheet structure, C-C=C stretching in Adenine and C—O and C—C stretching lipids and fatty acids are all in accordance as the results described in mean spectra. The positive loadings of the filtrate samples including 436 cm^{-1} (Lipids), 728 cm^{-1} (C—C stretching, proteins), 856 cm^{-1} (Amino acid), 1102 cm^{-1} (Phenylalanine (proteins)), 1255 cm^{-1} (Lipids), 1332 cm^{-1} (-C stretch of Phenyl and C_3—C_3 stretch/in-plane bend of proteins), 1450 cm^{-1} (CH_2 bending, CH functional groups in lipids, CH_2 bending (proteins)), 1590 cm^{-1} (Carbon particles), and 1658 cm^{-1} (Amide-I (proteins))are associated with the SERS spectral data sets of low viral load class of samples and show decrease in intensities as the viral load increases.

Fig. 5c(i) shows PCA scatter plot of centrifuged filtrate samples of medium viral load (11 samples) versus high viral load (8 samples) shown by green color and blue color dot, respectively. As shown in the plot these categories of low VL and medium VL are distinguished on the basis of principle component analysis. For medium vs high viral load, PC-1explains 46% of the variability in the two data sets and PC-2 explains 19%. The samples of the filtrates of medium viral load category are the clustered on the negative side and the samples of filtrates of high viral load are clustered on the positive side which shows that samples of medium and high viral load are well separated due to difference of biochemical changes reflected in the form of SERS spectral features due to changes in viral load. Fig. 5c(ii) shows the loadings of PCA of categories of low vs medium viral load samples. Remarkably, negative loadings are detected at 499 cm^{-1} (C—OH_3 torsion of methoxy), 639 cm^{-1} (amino acid), 817 cm^{-1} (C—C stretching collagen assignment), 888 cm^{-1} (protein band), 1008 cm^{-1} (phenylalanine), 1134 cm^{-1} (C—C stretching of protein), 1209 cm^{-1} (Amide III-Due to C—N stretching and N—H Bending), 1374 cm^{-1} (RNA base), 1411 cm^{-1} (lipid band), 1704 cm^{-1} (n(C_{55}O)OH (amino acids aspartic & glutamicacid) and 1774 cm^{-1} (lipid and fatty acid) which are associated the SERS spectral data set of filtrates of high VL, clustered in the negative axis. These features indicate that as the viral load increases, intensities of these features increase. Other peaks that are observed on the positive side of PC-1 loadings contain stretching of PO_4^{-3} phosphate, Amide I and Amide-III β-sheet structure, C-C=C stretching in Adenine and C—O and C—C stretching lipids and fatty acids are all in accordance as the results described in mean spectra. The positive loadings of the filtrate samples including 436 cm^{-1} (Lipids), 728 cm^{-1} (C—C stretching, proteins), 856 cm^{-1} (Amino acid), 1102 cm^{-1} (Phenylalanine (proteins)), 1255 cm^{-1} (Lipids), 1332 cm^{-1} (-C stretch of Phenyl and C_3—C_3 stretch/in-plane bend of proteins), 1450 cm^{-1} (CH_2 bending, CH functional groups in lipids, CH_2 bending (proteins)), 1590 cm^{-1} (Carbon particles), and 1658 cm^{-1} (Amide-I (proteins))are associated with the SERS spectral data sets of low viral load class of samples and show decrease in intensities as the viral load increases.

Fig. 6 shows pairwise PCA scatter plot and a combined loading of centrifuged filtrate of all viral load versus healthy samples are given. In Fig. 6(a) shows PCA scatter plot of centrifuged filtrate of low viral load (9 samples) versus healthy (13 samples) shown by blue color and pink color dots, respectively. As shown in the plot that categories of low VL and healthy samples are distinguished on the basis of principal component analysis. For low vs. healthy, PC-1 explains 50% of the variability in the two data sets and PC-2 explains 19%. The samples of filtrates of low viral load category are clustered on the positive side and the healthy samples are clustered on the negative side which shows that samples of Low VL and healthy samples are well separated due to difference of biochemical changes reflected in the form of SERS spectral features. Fig. 6(b) shows PCA scatter plot of centrifuged filtrate of medium viral load (9 samples) versus healthy (13 samples) shown by red color and pink color dots, respectively. As shown in the plot that categories of medium VL and healthy samples are distinguished on the basis of principal component analysis. For low vs. healthy, PC-1 explains 63% of the variability in the two data sets and PC-2 explains 17%. The samples of filtrates of medium viral load category are clustered on the positive side and the healthy samples are clustered on the negative side which shows that samples of medium VL and healthy samples are well separated due to difference of biochemical changes reflected in the form of SERS spectral features. Fig. 6(c) shows PCA scatter plot of centrifuged filtrate of high viral load (9 samples) versus healthy (13 samples) shown by green color and pink color dots, respectively. As shown in the plot that categories of high VL and healthy samples are distinguished on the basis of principal component analysis. For low vs. healthy, PC-1 explains 52% of the variability in the two data sets and PC-2 explains 22%. The samples of filtrates of high viral load category are clustered on the positive side and the healthy samples are clustered on the negative side which shows that samples of high VL and healthy samples are well separated due to difference of biochemical changes reflected in the form of SERS spectral features. Fig. 6(d) shows the loadings of PCA of categories of low vs. medium viral load samples. Remarkably, negative loadings are detected at 499 cm^{-1} (C—OH_3 torsion of methoxy), 639 cm^{-1} (amino acid), 817 cm^{-1} (C—C stretching collagen assignment), 888 cm^{-1} (protein band), 1008 cm^{-1} (phenylalanine), 1134 cm^{-1} (C—C stretching of protein), 1209 cm^{-1} (Amide III-Due to C—N stretching and N—H Bending), 1374 cm^{-1} (RNA base), 1411 cm^{-1} (lipid band), 1704 cm^{-1} (n(C_{55}O)OH (amino acids aspartic & glutamicacid) and 1774 cm^{-1} (lipid and fatty acid) which are associated the SERS spectral data set of filtrates of high VL, clustered in the negative axis. These features indicate that as the viral load increases, intensities of these features increase. Other peaks that are observed on the positive side of PC-1 loadings contain stretching of PO_4^{-3} phosphate, Amide I and Amide-III β-sheet structure, C-C=C stretching in Adenine and C—O and C—C stretching lipids and fatty acids are all in accordance as the results described in

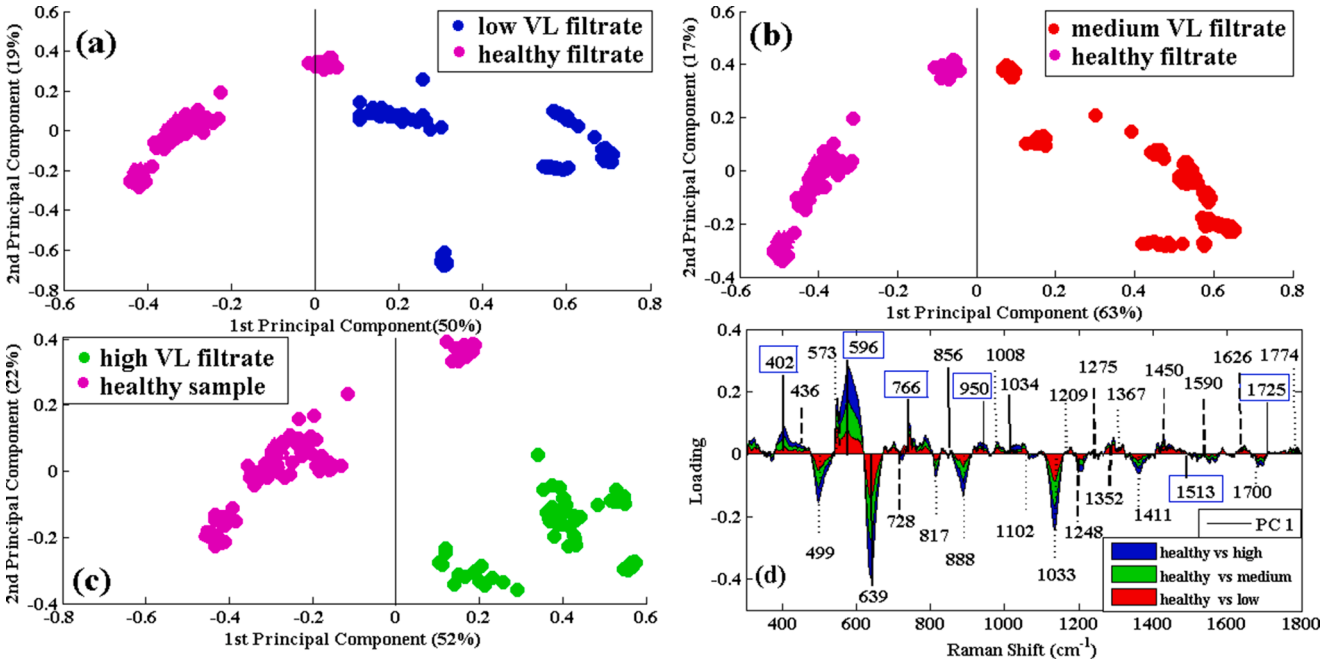


Fig. 6. Pairwise PCA scatter plots of SERS spectral data sets of filtrate portions of healthy and HCV positive serum samples including low, medium and high viral loads (a) healthy versus Low viral load filtrate (b) healthy versus medium viral load filtrate (c) healthy versus high viral load filtrate and (d) PCA loadings of low, medium and high viral load classes versus healthy samples.

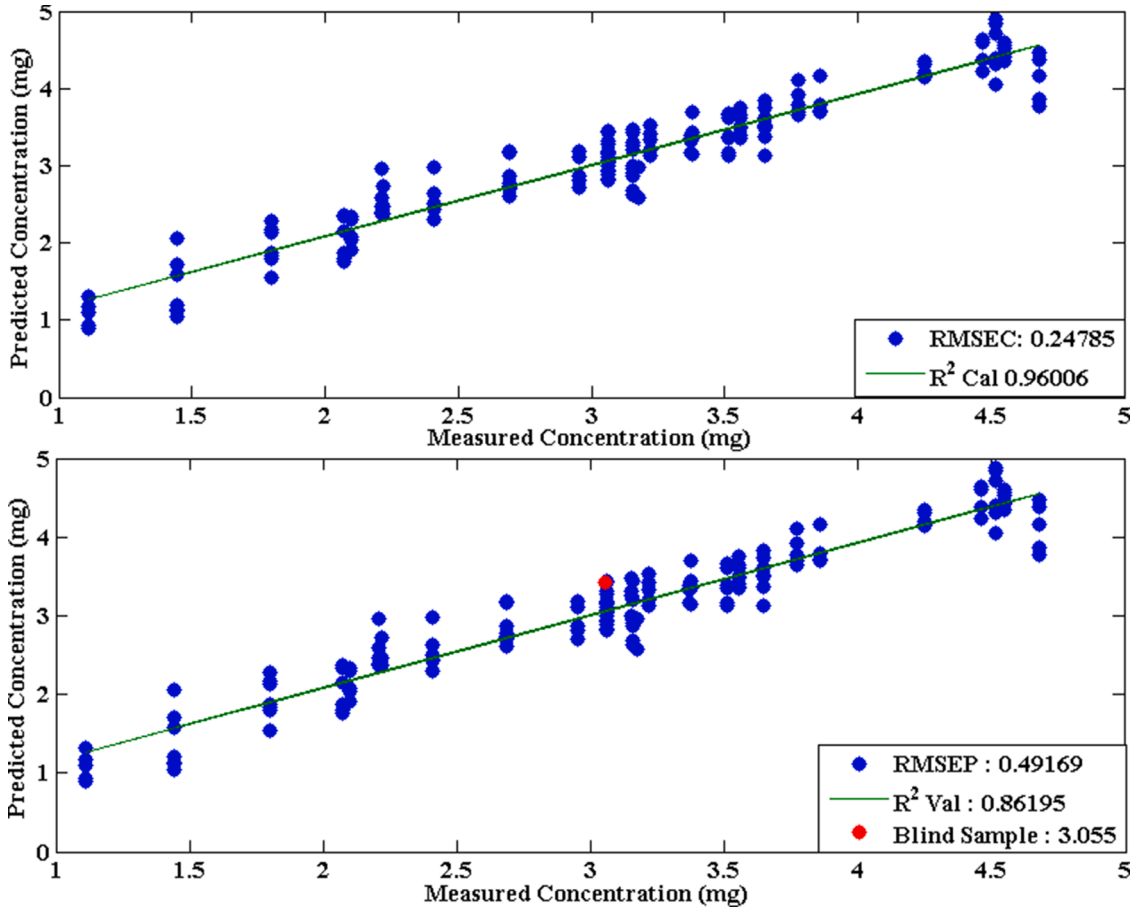


Fig. 7. PLSR model of for SERS spectral datasets of filtrate portions of HCV positive serum samples; calibration model (a) and prediction model (validation model) for prediction of viral load of blind/unknown sample.

mean spectra. The positive loadings of the filtrate samples including 436 cm^{-1} (Lipids), 728 cm^{-1} (C—C stretching, proteins), 856 cm^{-1} (Amino acid)' 1102 cm^{-1} (Phenylalanine (proteins)), 1255 cm^{-1} (Lipids), 1332 cm^{-1} (-C stretch of Phenyl and C_3—C_3 stretch/in-plane bend of proteins), 1450 cm^{-1} (CH_2 bending, CH functional groups in lipids, CH_2 bending (proteins)), 1590 cm^{-1} (Carbon particles), and 1658 cm^{-1} (Amide-I (proteins))are associated with the SERS spectral data sets of low viral load class of samples and show decrease in intensities as the viral load increases.

3.3. Partial least square regression analysis

PLSR analysis is applied to evaluate the capability of Surface Enhance Raman Spectroscopy for the prediction of viral loads of all the filtrate samples of HCV samples. SERS spectral data of all the samples (420 spectra) are put in a single matrix and randomly divided into two halves, 60% (255 spectra) and 40% (165 spectra), used as calibration dataset and validation dataset/test dataset, respectively, for the development of the PLSR model and to prevent dataset bias. Notably, leave one sample (15 spectra of that sample) out cross validation is employed in order to avoid overfitting of the model. Thirteen latent variables are found optimum for the development of PLSR model of 43 filtrate samples presented in Table 1.

Fig. 7(a) displays the calibration of PLSR model and plot (b) is the predicted PLSR model. The calculated value of goodness (R^2 calculated) for calibration is 0.96 and for prediction is 0.86 which indicates the validity of the regression model. In order to further elaborate the validity of our developed PLSR model, one sample was considered as blind sample and PLS regression tool was applied for unknown blind sample S19. The predicted value of blind sample S19 was found to be 3.055 as shown in Fig. 7(b).

4. Conclusions

Surface-enhanced Raman spectroscopy (SERS) has been used for the analysis of filtrate portion of HCV positive serum samples of different levels of viral loads and healthy persons. Different SERS spectral features are identified which are associated with the HCV infection and also with the increasing viral load values. For this purpose, centrifugal filtration technique is found very helpful in order to concentrate the HCV related proteins by removing proteins bigger in size than 50 KDa. Furthermore, Principal Component Analysis (PCA) of the SERS data indicated a significant differentiation between filtrate samples of different viral loads. Biochemical differences found between infected and healthy serum samples were used to create the PLSR model for predicting HCV viral load with 99% accuracy and reliability (R2) of 0.96. Finally, the method presented here has potential to be used as a qualitative and quantitative technique for the diagnosis of hepatitis C.

Funding information

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

CRediT authorship contribution statement

Samra Shakeel: Writing – original draft. **Haq Nawaz:** Conceptualization. **Muhammad Irfan Majeed:** Project administration. **Nosheen Rashid:** Investigation. **Muhammad Rizwan Javed:** Methodology. **Ayesha Tariq:** Data curation. **Beenish Majeed:** Data curation. **Aafia Sehar:** Formal analysis. **Sania Murtaza:** Formal analysis. **Nimra Sadaf:** Software. **Gull Rimsha:** Software. **Imran Amin:** Resources.

Declaration of Competing Interest

The authors have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] C.W. Shepard, L. Finelli, M.J. Alter, Global epidemiology of hepatitis C virus infection, Lancet Infect. Dis. 5 (9) (2005) 558–567.
[2] D. Lavanchy, The global burden of hepatitis C, Liver Int. 29 (2009) 74–81.
[3] N. Ford, et al., Chronic hepatitis C treatment outcomes in low-and middle-income countries: a systematic review and meta-analysis, Bull. World Health Organ. 90 (2012) 540–550.
[4] S.L. Chen, T.R. Morgan, The natural history of hepatitis C virus (HCV) infection, Int. J. Med. Sci. 3 (2) (2006) 47.
[5] B. Lange, et al., Diagnostic accuracy of serological diagnosis of hepatitis C and B using dried blood spot samples (DBS): two systematic reviews and meta-analyses, BMC Infect. Dis. 17 (1) (2017) 87–106.
[6] S.H.H. Ahmed, et al., Evaluation of some available HCV antibody detection tests (ELISA, chemiluminescence, immune assay) and RT-PCR assay in the diagnosis of Hepatitis C virus infection, Egypt. J. Hosp. Med. 72 (7) (2018) 4874–4879.
[7] M. Swellam, M.S. Mahmoud, A.A.F. Ali, Diagnosis of hepatitis C virus infection by enzyme-linked immunosorbent assay and reverse transcriptase-nested polymerase chain reaction: a comparative evaluation, IUBMB Life 63 (6) (2011) 430–434.
[8] T. Mahmood, et al., Raman spectral analysis for rapid screening of dengue infection, Spectrochim. Acta Part A 200 (2018) 136–142.
[9] C.J. Smith, A.M. Osborn, Advantages and limitations of quantitative PCR (Q-PCR)-based approaches in microbial ecology, FEMS Microbiol. Ecol. 67 (1) (2009) 6–20.
[10] M. Enomoto, et al., Comparison of real-time quantitative polymerase chain reaction with three other assays for quantitation of hepatitis C virus, J. Gastroenterol. Hepatol. 16 (8) (2001) 904–909.
[11] A.M. Ranieri, et al., Complementary approaches to imaging subcellular lipid architectures in live bacteria using phosphorescent iridium complexes and Raman spectroscopy, Chem. Eur. J. 25 (45) (2019) 10566–10570.
[12] M.H. Vos, T. Scopigno, Ultrafast dynamics and vibrational relaxation in six-coordinate heme proteins revealed by femtosecond stimulated Raman, J. Am. Chem. Soc. 142 (2020) 2285–2292.
[13] L. Ashton, et al., UV resonance Raman spectroscopy: a process analytical tool for host cell DNA and RNA dynamics in mammalian cell lines, J. Chem. Technol. Biotechnol. 90 (2) (2015) 237–243.
[14] C.K. Yao, et al., Spatially reinforced nano-cavity array as the SERS-active substrate for detecting hepatitis virus core antigen at low concentrations, Sens. Actuators B 174 (2012) 478–484.
[15] S. Herrick, et al., Fibrinogen, Int. J. Biochem. Cell Biol. 31 (7) (1999) 741–746.
[16] S. Ahmad, et al., Characterization and prediction of viral loads of hepatitis B serum samples by using surface-enhance Raman spectroscopy (SERS), Photodiagn. Photodyn. Ther. (2021), 102386.
[17] M. Kashif, et al., Surface enhanced Raman spectroscopy of the serum samples for the diagnosis of hepatitis C and prediction of the viral loads, Spectrochim. Acta Part A 242 (2020), 118729.
[18] S. Chevaliez, Virological tools to diagnose and monitor hepatitis C virus infection, Clin. Microbiol. Infect. 17 (2) (2011) 116–121.
[19] R.Z.A. Bari, et al., Surface-enhanced Raman spectroscopic analysis of centrifugally filtered HBV serum samples, Photodiagn. Photodyn. Ther. (2022), 102808.
[20] L. Shao, et al., Fast and non-invasive serum detection technology based on surface-enhanced Raman spectroscopy and multivariate statistical analysis for liver disease, Nanomed. Nanotechnol. Biol. Med. 14 (2) (2018) 451–459.
[21] F. Bonnier, et al., Ultra-filtration of human serum for improved quantitative analysis of low molecular weight biomarkers using ATR-IR spectroscopy, Analyst 142 (8) (2017) 1285–1298.
[22] N. Stone, et al., Raman spectroscopy for identification of epithelial cancers, Faraday Discuss. 126 (2004) 141–157.
[23] C. Krafft, et al., Near infrared Raman spectra of human brain lipids, Spectrochim. Acta Part A 61 (7) (2005) 1529–1535.
[24] H. Schulz, M. Baranska, Identification and quantification of valuable plant substances by IR and Raman spectroscopy, Vib. Spectrosc. 43 (1) (2007) 13–25.
[25] J.W. Chan, et al., Micro-Raman spectroscopy detects individual neoplastic and normal hematopoietic cells, Biophys. J. 90 (2) (2006) 648–656.
[26] C.J. Frank, R.L. McCreery, D.C. Redd, Raman spectroscopy of normal and diseased human breast tissues, Anal. Chem. 67 (5) (1995) 777–783.
[27] S. Farquharson, et al., Analysis of 5-fluorouracil in saliva using surface-enhanced Raman spectroscopy, J. Raman Spectrosc. 36 (3) (2005) 208–212. Int. J. Orig. Work Aspects of Raman Spectroscopy, Including Higher Order Processes, and also Brillouin and Rayleigh Scattering.
[28] W.T. Cheng, et al., Micro-Raman spectroscopy used to identify and grade human skin pilomatrixoma, Microsc. Res. Tech. 68 (2) (2005) 75–79.
[29] E. Hanlon, et al., Prospects for *in vivo* Raman spectroscopy, Phys. Med. Biol. 45 (2) (2000) R1.
[30] Z. Movasaghi, S. Rehman, I.U. Rehman, Raman spectroscopy of biological tissues, Appl. Spectrosc. Rev. 42 (5) (2007) 493–541.
[31] J. Laska, J. Widlarz, Spectroscopic and structural characterization of low molecular weight fractions of polyaniline, Polymer 46 (5) (2005) 1485–1495.
[32] D. Naumann, Infrared and NIR Raman spectroscopy in medical microbiology, Int. Soc. Opt. Photonics (1998).
[33] G. Shetty, et al., Raman spectroscopy: elucidation of biochemical changes in carcinogenesis of oesophagus, Br. J. Cancer 94 (10) (2006) 1460–1464.

[34] A. Ruiz-Chica, et al., Characterization by Raman spectroscopy of conformational changes on guanine–cytosine and adenine–thymine oligonucleotides induced by aminooxy analogues of spermidine, J. Raman Spectrosc. 35 (2) (2004) 93–100.
[35] Y.K. Min, et al., 1064nm near-infrared multichannel Raman spectroscopy of fresh human lung tissues, J. Raman Spectrosc. 36 (1) (2005) 73–76. An International Journal for Original Work in all Aspects of Raman Spectroscopy, Including Higher Order Processes, and also Brillouin and Rayleigh Scattering.
[36] R.J. Lakshmi, et al., Tissue Raman spectroscopy for the study of radiation damage: brain irradiation of mice, Radiat. Res. 157 (2) (2002) 175–182.
[37] A. Mahadevan-Jansen, R. Richards-Kortum, Raman spectroscopy for cancer detection: a review, in: Proceedings of the 19th Annual International Conference of the IEEE Engineering in Medicine and Biology Society.'Magnificent Milestones and Emerging Opportunities in Medical Engineering'(Cat. No. 97CH36136), IEEE, 1997.
[38] N.J. Kline, P.J. Treado, Raman chemical imaging of breast tissue, J. Raman Spectrosc. 28 (2–3) (1997) 119–124.
[39] A.A. Makki, et al., Qualitative and quantitative analysis of therapeutic solutions using Raman and infrared spectroscopy, Spectrochim. Acta Part A 218 (2019) 97–108.
[40] J. Saade, et al., Identification of hepatitis C in human blood serum by near-infrared Raman spectroscopy, Spectroscopy 22 (5) (2008) 387–395.
[41] A.A. Butt, et al., Changes in circulating lipids level over time after acquiring HCV infection: results from ERCHIVES, BMC Infect. Dis. 15 (1) (2015) 1–9.