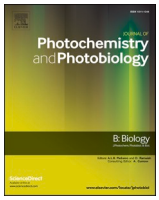
Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Journal of Photochemistry & Photobiology, B: Biology 234 (2022) 112545

Contents lists available at ScienceDirect

Journal of Photochemistry & Photobiology, B: Biology

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



A non-invasive ultrasensitive diagnostic approach for COVID-19 infection using salivary label-free SERS fingerprinting and artificial intelligence

Varsha Karunakaran [a,1,2], Manu M. Joseph [a,1,2], Induprabha Yadev [b], Himanshu Sharma [c,d], Kottarathil Shamna [a], Sumeet Saurav [c,d], Remanan Pushpa Sreejith [a,d], Veena Anand [b], Rosenara Beegum [b], S. Regi David [b], Thomas Iype [b], K.L. Sarada Devi [b], A. Nizarudheen [b], M.S. Sharmad [b], Rishi Sharma [c,d], Ravindra Mukhiya [c,d], Eshwar Thouti [c,d], Karuvath Yoosaf [a,d], Joshy Joseph [a,d], P. Sujatha Devi [a,d], S. Savithri [a,d], Ajay Agarwal [e], Sanjay Singh [c,d,*], Kaustabh Kumar Maiti [a,d,*]

[a] CSIR – National Institute for Interdisciplinary Science and Technology, Thiruvananthapuram, Kerala, India
[b] Government Medical College, Thiruvananthapuram, Kerala, India
[c] CSIR-Central Electronics Engineering Research Institute, Pilani 333031, India.
[d] Academy of Scientific and Innovative Research (AcSIR), Ghaziabad 201002, India
[e] Department of Electrical Engineering, Indian Institute of Technology Jodhpur, Rajasthan 342037, India

ARTICLE INFO

Keywords:
Surface enhanced Raman spectroscopy
Saliva
COVID-19
Label-free
Diagnosis
Artificial intelligence

ABSTRACT

Clinical diagnostics for SARS-CoV-2 infection usually comprises the sampling of throat or nasopharyngeal swabs that are invasive and create patient discomfort. Hence, saliva is attempted as a sample of choice for the management of COVID-19 outbreaks that cripples the global healthcare system. Although limited by the risk of eliciting false-negative and positive results, tedious test procedures, requirement of specialized laboratories, and expensive reagents, nucleic acid-based tests remain the gold standard for COVID-19 diagnostics. However, genetic diversity of the virus due to rapid mutations limits the efficiency of nucleic acid-based tests. Herein, we have demonstrated the simplest screening modality based on label-free surface enhanced Raman scattering (LF-SERS) for scrutinizing the SARS-CoV-2-mediated molecular-level changes of the saliva samples among healthy, COVID-19 infected and COVID-19 recovered subjects. Moreover, our LF-SERS technique enabled to differentiate the three classes of corona virus spike protein derived from SARS-CoV-2, SARS-CoV and MERS-CoV. Raman spectral data was further decoded, segregated and effectively managed with the aid of machine learning algorithms. The classification models built upon biochemical signature-based discrimination method of the COVID-19 condition from the patient saliva ensured high accuracy, specificity, and sensitivity. The trained support vector machine (SVM) classifier achieved a prediction accuracy of 95% and F1-score of 94.73%, and 95.28% for healthy and COVID-19 infected patients respectively. The current approach not only differentiate SARS-CoV-2 infection with healthy controls but also predicted a distinct fingerprint for different stages of patient recovery. Employing portable hand-held Raman spectrophotometer as the instrument and saliva as the sample of choice will guarantee a rapid and non-invasive diagnostic strategy to warrant or assure patient comfort and large-scale population screening for SARS-CoV-2 infection and monitoring the recovery process.

Abbreviations: SERS, Surface enhanced Raman spectroscopy; AuNPs, Gold nanoparticles; PCA, Principal Component Analysis; SVM, Support Vector Machine.

* Corresponding authors at: Academy of Scientific and Innovative Research (AcSIR), Ghaziabad 201002, India.
E-mail addresses: sanjay@ceeri.res.in (S. Singh), kkmaiti@niist.res.in (K.K. Maiti).

[1] Both authors are having equal contribution.

[2] Present Addresses: CSIR-National Institute for Interdisciplinary Science & Technology (NIIST), Chemical Sciences &Technology Division (CSTD), Industrial Estate, Thiruvananthapuram 695019, Kerala, India.

https://doi.org/10.1016/j.jphotobiol.2022.112545
Received 11 February 2022; Received in revised form 27 July 2022; Accepted 16 August 2022
Available online 19 August 2022

1. Introduction

Viruses are profuse biological entities capable of infecting every form of life. Coronaviruses (CoVs) frequently cause respiratory illness in humans. Middle East respiratory syndrome coronavirus (MERS-CoV) and severe acute respiratory syndrome coronavirus (SARS-CoV) are two extremely transmissible and pathogenic viruses that infected humans at the outset of the 21st century. The global outbreak of the 2019 novel coronavirus (2019-nCoV), triggered by SARS-CoV-2 brings about the announcement of a pandemic by the World Health Organization (WHO). Till now, SARS-CoV-2 has infected more than 266 million people and has claimed over 5.26 million lives around the globe. SARS-CoV-2 has a single-stranded positive-sense RNA genome and further crowned with a protein-spiked envelope. The structural proteins encompass spike surface glycoprotein (S), matrix protein (M), a small envelope protein (E), and nucleocapsid protein (N). Spike protein of SARS-CoV recognizes host angiotensin-converting enzyme 2(ACE2) for the cell entry and subsequent replication [1,2]. SARS-CoV-2 principally deploys through the respiratory tract secretions, droplets, and by direct contact. Reverse transcription-polymerase chain reaction (RT-PCR) and reverse-transcription loop-mediated isothermal amplification (RT-LAMP) are the gold standard nucleic acid-based diagnostic tests for COVID-19. These tests are associated with the occurrence of false-negative and positive results, longer analytical protocol, requirement of specialized laboratories and expensive reagents. Genetic diversity caused by rapid mutations in the loci recognized by DNA primers may potentially hinder sensitivity and selectivity of nucleic acid-based COVID-19 diagnosis [3]. Moreover, the U.S. Food and Drug Administration (FDA) have alerted that the mutant strains may severely compromise the reliability of nucleic acid-based molecular diagnosis [4]. Rapid antigen and antibody based serological tests are invasive which consists of lateral flow-based hand-held disposable strips and high throughput enzyme-linked immunosorbent assay but the sensitivity and reliability are far behind the nucleic acid-based diagnostics [5].

Currently employed COVID-19 diagnosis uses nasopharyngeal swabs wherein the sample collection procedure is invasive, risky, cause discomfort to the patients and occasionally induces coughing and bleeding. Since the spread of the virus occurs through respiratory droplets and aerosols, saliva was suggested as a substitute for nasopharyngeal swabs. Saliva can be collected without any invasive procedures and with proper guidelines, patients can collect saliva themselves to minimize the risk of virus transmission to healthcare personnel. Saliva displayed a high concordance rate with nasopharyngeal specimens in the detection of viruses, including coronaviruses. The salivary SARS-CoV-2 load will reach a peak during the first week from symptom onset and could be detected as long as 20 days, suggesting its use for monitoring viral clearance [6]. Moreover, the presence of SARS-CoV-specific secretory IgA and IgG responses was reported in saliva [7]. SARS-CoV-2 diagnosis using saliva samples could especially help countries with a high population density to conduct mass population screening and effectively mitigate COVID-19 spread.

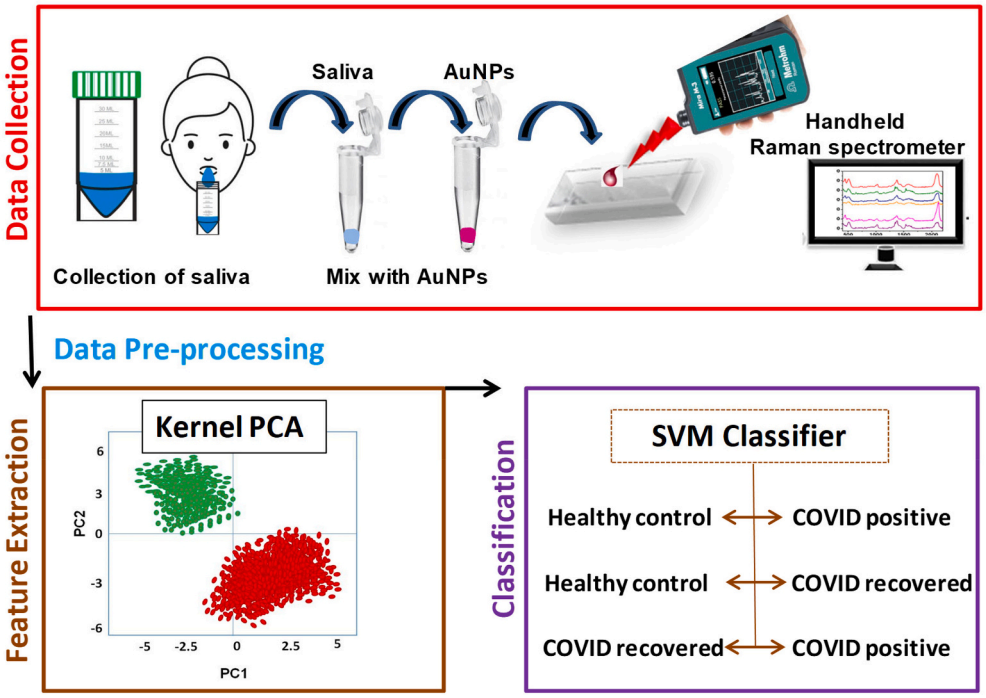
Raman spectroscopy is a useful tool for scrutinizing the molecular-level changes in the biological specimen by making use of the inelastically scattered light for the recognition of vibrational states of the biomolecules. In order to amplify the inherently weak Raman scattering, surface-enhanced Raman scattering (SERS) has emerged, which enhances signal intensity of molecular vibration up to a magnitude of 10^8–10^{14} folds with the aid of metallic nanoparticles [8]. On the grounds of facile synthesis, tunable size and high-localized surface plasmon resonance, gold nanoparticles (AuNPs) are largely employed as SERS substrates. Several attempts were made using the SERS platform for the rapid identification of pathogens including different forms of viruses. The virus envelope is made up of a unique set of proteins and lipids and hence can generate distinct Raman fingerprints. Recently, many SERS-based platforms were investigated for SARS-CoV-2 detection with high sensitivity but most of the approaches were accompanied with the drawback of complexity in the design of probes, lengthy processes and the requirement of invasive sample collection procedures [9–11]. Raman spectral data abstracted from the label-free analysis of biomolecules can be decoded, segregated, and effectively managed with the aid of Machine Learning (ML) and Deep Learning (DL) algorithms. Computer-aided chemometric approach for the design of classification and prediction models with the biological data of spectroscopic studies have previously been established for a range of applications like disease diagnosis and prognosis, proteomics, genomics, metabolomics and evaluation of therapy [12].

The purpose of the current investigation was to explore the feasibility of employing SERS as a diagnostic modality for the label-free discrimination of SARS-CoV-2 infection from saliva samples. Portable hand-held Raman spectrophotometer has been utilized to secure the Raman fingerprints from the saliva samples where colloidal AuNPs was mixed with saliva in order to evaluate the enhanced Raman peaks. The classification models built-up by machine learning and deep learning algorithms based on salivary SERS spectra that will ensure high accuracy, specificity, and sensitivity to the approach. Considering the significant damage created by several waves of COVID viral transmission and the prediction for future outbreaks, an alternative strategy for diagnosis in a non-invasive manner with high accuracy is demanded to ensure mass population screening to effectually tame the next wave. Here, we aim to develop a portable, fast, and highly sensitive technique for the identification of COVID-19 infection by label-free surface enhanced Raman spectral analysis and principal component analysis (PCA) for the identification of unique bio-molecular Raman fingerprints of COVID-19 infection in the body fluids like saliva and classify the healthy and COVID-19 infected samples using trained support vector machine (SVM) classifier. Hence, we believe that our current approach (Scheme 1) will certainly be translated to the bedside application after a large cohort study in a time-bound manner.

2. Experimental Methods

Synthesis and characterization of gold nanoparticles. Gold nanoparticles (AuNPs) were synthesized by the citrate reduction method as described previously [13]. All glassware were cleaned with aqua regia, then rinsed and steeped in Milli-Q water before use. Briefly, MilliQ water (50 mL) was heated until boiling to which 50 μL gold chloride (0.25 M) was added. After 10 min of boiling, 125 μL of trisodium citrate solution (0.1 M) was added which resulted in a color change from purple to red in 5 min. The solution was allowed to cool for about 90 min with constant stirring. The as-prepared nanoparticles were concentrated by centrifugation for 45 min at 4000 rpm for SERS analysis. Later the size, shape, and monodispersity of the prepared AuNPs were confirmed by ultraviolet-visible (UV–Vis) spectroscopy, dynamic light scattering (DLS), and high-resolution transmission electron microscopy (HR-TEM).

Optimization of SERS studies. SERS experiments were performed with the aid of a handheld Raman spectroscope (Metrohm, USA) with a laser beam of 785 nm at 6 mW power directed to the sample through collection optics termed universal attachment (6.07506.010) with a numerical aperture (NA = 0.50 mm working distance and 0.2–2.5 mm spot size). Position 2 in the universal attachment was used for the study which is approximately 3 mm from the end of the attachment. The spectral resolution was 8–10 cm^{-1} (FWHM) with beam divergence 2°. The detection technique was Orbital Raster Scan (ORS) to average over the sample. The laser class was 3B according to EN 60825–1. Different parameters like laser power, number of spectra required, accumulation time, sample positioning, sample volume required, and sample to AuNPs ratio were optimized with multiple experiments. The spectral data were analyzed using Miracal software. Briefly, before each analysis, the instrument was calibrated using the reference Raman reporter MBA (4-Mercapto benzoic acid, 100 μM) and the intensity values were noted down. Spectra were acquired in the region between 400 and 2300 cm^{-1}. While mentioning about the problems associated with the



Scheme 1. Schematic representation for workflow employed for the proposed strategy aiming COVID-19 diagnostics using SERS and AI with saliva as the sample of choice.

instrumentation and sample handling, impurities and inhibitors within saliva may be problematic to get expected results. Further we observed that high content of localized mucus particles can produce misleading Raman spectra and hence these suspended particles need to be carefully avoided before mixing with AuNPs and subsequent Raman studies. Moreover, presence of high lux flashlights or high exposure to sunlight during the SERS spectral acquisition was found to be problematic with the generation of un warrant Raman peaks. Hence the sample acquisition procedure needs to be performed in a less-light / dark room for an interference-free Raman spectrum.

Sample collection. The study was approved and conducted following the directives of the ethical committee institutional review board at the Government Medical College, Thiruvananthapuram, Kerala, India (Human Ethics Committee approval no: HEC No.05/19/2020/MCT). For initial studies and optimization, bovine serum albumin (BSA) was used for spiking with health saliva samples. Later, S1 spike protein of SARS-CoV-2 (Sino Biological, Inc., Cat no: 40591-V08H), SARS-CoV (Sino Biological, Inc., Cat No: 40150-V08B1), and MERS-CoV (Sino Biological, Inc., Cat no: 40069-V08B1) was employed. For the studies with COVID patients, saliva samples from clinically confirmed COVID-19 patients were obtained after getting informed consent. The saliva samples were collected after 3 rounds of mouthwash by the drooling method in a sterile 50 mL centrifuge tube and stored at 4 °C till use. A saliva drop (10 μL) was mixed with AuNPs (40 μL) deposited on a glass slide covered with commercially available aluminum foil and subjected to Raman spectral acquisition. Information related to the sample size, number of subjects used, number of spectra taken from each subject, total number of spectra used for the studies, variance ratio of the principal components, spectral range (cm^{-1}) used and facts related to the training and testing data set employed for different studies using viral proteins and saliva samples are illustrated in S1-S7 (Supporting information).

Initial pre-processing of data. The raw spectral data were pre-processed by Miracal software before the statistical analysis in order to remove the interference noises, cosmic rays, and oversaturated spectra. The background was removed by the 4th degree polynomial function and smoothening of the SERS spectra was done by the Savitzky-Golay smoothing which in turn was normalized in the region of 400–2300 cm^{-1}. The mean spectrum of each group of the pre-processed data was then determined using the Origin Pro 8.0 software. The comparison between the spectra of the control saliva and the treated group was made through the subtraction of different mean spectra and the shifts of the different peaks in the subtracted spectra were assigned to the molecular structures and biochemical components based on the literature reports.

Baseline correction of data. It was observed that the spectral data received after the initial pre-processing step contains the baseline drift and in order to correct the baseline drift of the spectral data, a novel algorithm named asymmetrically reweighted penalized least squares smoothing (ARPLS) [14] was used. The above algorithm uses lambda as its input parameter for the smoothness of the spectral data. The greater the value of lambda the smoother the output spectra. By the repeated experiments we have found that the lambda value at 10^6 provides the smoother baseline corrected spectral curve with lowest root mean square error (RMSE) value. Thus, for all our experiments we have used the lambda value of 10 [6].

Prominent feature selection and machine learning algorithm. The handheld Raman spectrometer provides 1650 data points in the range of 400–2300 cm^{-1}. We have taken data points in the range of 500 to 2150 cm^{-1} as inputs for our machine-learning framework (feature selection and classification). The baseline-corrected data was taken as input into the kernel principal component analysis (kernel PCA) [15] algorithm for dimensionality reduction and prominent feature selection where Radial basis function (Rbf) is used as the kernel. The extracted features by kernel PCA were used to train the support vector machine (SVM) classifier. Details of the classifier training and testing are mentioned below.

Classification of samples. The complete process of classifying the Raman spectral data into healthy and COVID-19 infected categories consists of three major steps: (i) feature extraction using kernel PCA (ii) training of SVM classifier, and (iii) testing of the trained SVM classifier. To extract the features from the baseline corrected Raman spectral data, kernel PCA was used on the whole dataset. Following the feature extraction-using kernel PCA on the whole dataset, the dataset was randomly divided into training data and testing data in the ratio of

75:25. To design the classifier we have used the libSVM library [15]. The SVM classifier was trained using 75% of the data. After completion of the training, the trained SVM classifier was tested on the remaining 25% of the data. In the case of the SVM classifier, the degrees of freedom are in fact the number of training samples in the dataset. Each training sample can become a support vector and thus can contribute to the separating plane in case of binary SVM classifier and separating hyperplane in case of the multi-class SVM classifier. One can use the degree of freedom of the model to predict overfitting. When we have more parameters (degree of freedom) than observations, there is a risk of overfitting the training dataset. This is intuitive, as parameters in the model than the observations may configure the model to predict the training data correctly and exactly, which usually indicates overfitting. The accuracy of the trained SVM classifier with respect to each of our dataset is reported in the respective sub-sections of the manuscript.

3. Results and Discussion

3.1. Optimization of SERS Fingerprinting Using AuNPs

SERS substrates are the key players behind the generation of hotspots to amplify the inherently weak Raman signals of analytes or the reporter molecules. AuNPs are the most exploited SERS substrates to amplify the Raman signatures of the biomolecules that would otherwise stay untraceable due to their lower Raman cross-sections. Spherical AuNPs within the size range of 40–45 nm are reported to be excellent for label-free Raman fingerprinting of biological analytes [16]. AuNPs were prepared, characterized (Fig. S1 a-c) and utilized as a SERS substrate for tracing biomolecular changes in the Raman signature. For evaluating the Raman enhancement effect, 4-mercaptobenzoic acid (MBA) was used as a reference standard and the instrumental parameters were optimized. The signature SERS peaks of MBA for a range of concentrations were recorded to reveal the concentration dependent intensity variations of the analyte molecule (Fig. 1a, S2a). The peak at 1084 cm^{-1} corresponds to v8a (a1) aromatic ring vibration having the C—S stretching mode and the one at 1586 cm^{-1} indicates v12 (a1) aromatic ring breathing mode of the MBA molecule [17]. With the portable handheld Raman spectroscope of a 785 nm laser and a power of 6 mW, an integration time of 10s and 3 accumulations with appropriate focal length yielded optimal intensity with well-resolved Raman peaks of MBA molecules. The high signal enhancement of AuNPs enabled future exploration in a SERS platform. The appreciable spectral reproducibility could be due to the optimum size, surface charge, and stable nature of AuNPs. We have recently illustrated that the size, shape, and surface charge of AuNPs are key parameters regulating their interaction with the biological system to influence the capacity to be used for SERS studies [18].

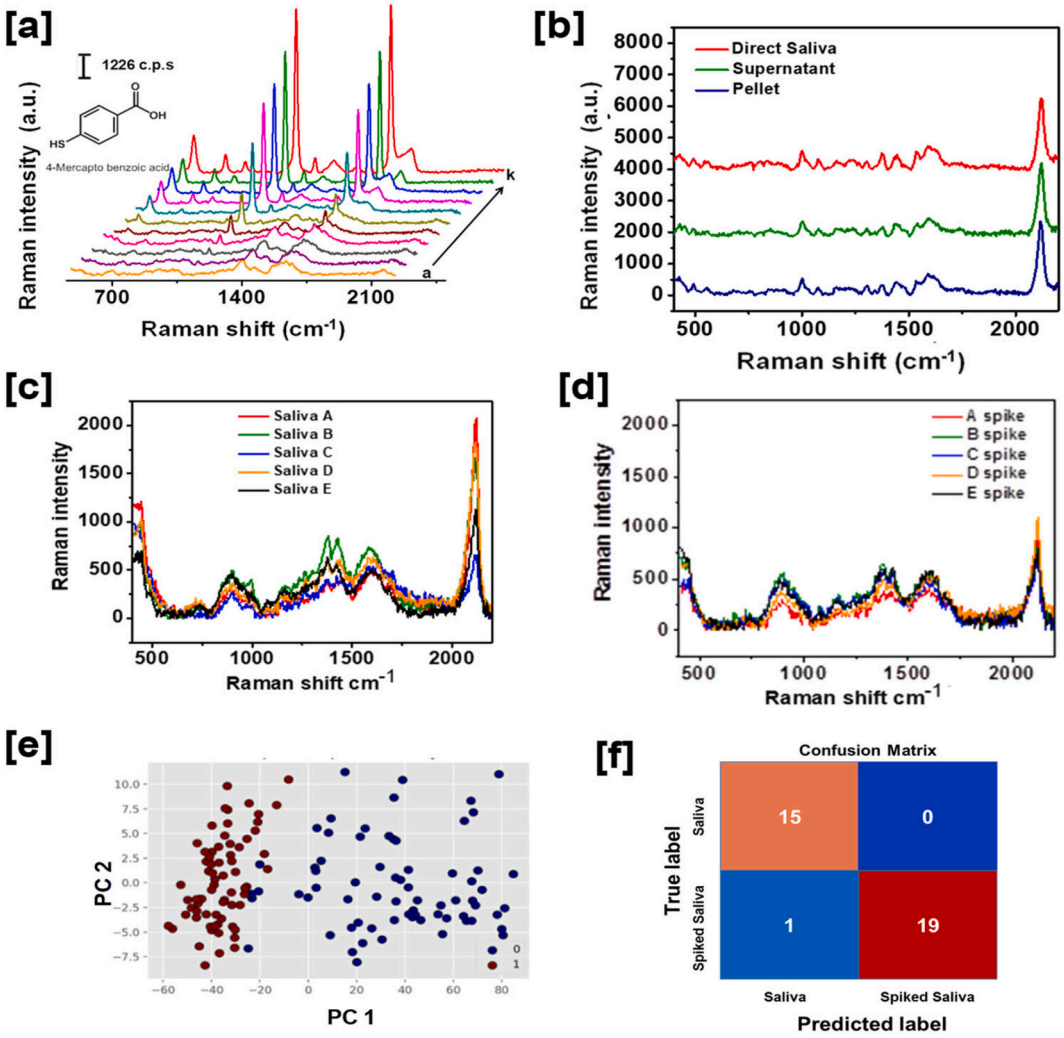


Fig. 1. [a] Waterfall SERS spectrum of MBA with different concentrations a) 0.97, b) 1.9, c) 3.9, d) 7.8, e) 15.6, f) 31.2, g) 62.5, h) 125, i) 250, j) 500 and k) 1000 μM, the insert figure shows the structure of MBA. [b] Stacked SERS spectrum of direct saliva, saliva supernatant, and saliva pellet for the process of optimization of sample of choice. Raman spectra of [c] saliva taken from different subjects and [d] spiked with model protein BSA (1 μg/mL). [e] PCA and [f] SVM analysis.

3.2. SERS Spectral Evaluation of Saliva Samples

Saliva samples were collected voluntarily by the healthy subjects using the drooling method. Each sample was separated into [1] saliva as such (direct saliva), [2] pellet (obtained after centrifugation of the sample at 3000 rpm for 5 min), and [3] supernatant (obtained from the previous step of centrifugation). The samples were subjected to Raman spectral analysis with the pre-optimized instrument settings and experimental conditions. Well-resolved Raman spectra were obtained from all the three types of samples of each subject (Fig. 1b). Although minor subject-to-subject variations were visible, the presented spectra were mainly dominated by the peaks attributed to C—N stretching and CH_3 vibration in protein backbone (1156 and 1303 cm^{-1}), cholesterol and ester at 425 cm^{-1}, glycogen at 490 cm^{-1}, and by the signal at 678, 729, 826 and 1373 cm^{-1} assigned to the phosphate backbone and nitrogenous bases from stretching of nucleic acids. Further, the peak at 2126 cm^{-1} represents the C—N— thiocyanate peak which is a significant fingerprint from the saliva [8,19,20] (Table S1). Notable biochemical Raman bands in the SERS spectra can tentatively be assigned to the various components in the saliva. Since we could not figure out any major variation between the three samples of each subject, the spectra were subjected to chemometric analysis using PCA. There was no significant difference in the Raman spectral pattern between saliva as such, pellet or supernatant of every subject (Fig. S2b, c). These observations prompted us to use saliva as such, without any sample processing procedures for all the subsequent studies. Our observation is in concordant with multiple recently published articles supporting saliva as an alternative sample for SERS analysis [21].

3.3. Spiking Studies with Saliva

The current way of employing SERS for abstracting molecular fingerprints from saliva is advantageous because of the non-invasive nature of sample collection protocol and rapid testing procedure. We next performed spiking studies on saliva by choosing a model protein bovine serum albumin (BSA) with clinically relevant concentration. It was observed that even with a lower concentration (1 μg/mL), separation was distinct between spiked and non-spiked saliva of all the subjects under investigation (Fig. 1c, d; S3a-c and Table S2). Further, the optimization was extended with a sample size of 15 subjects wherein, a visually distinguishable pattern difference was observed. Two components with variance score of kernel PCA and LDA (Fig. 1e, S3d) were able to extract the prominent features from the 14 subjects and a binary SVM classifier was trained on the features extracted from kernel PCA. Evaluated on the test samples, the trained SVM classifier achieved a F1-score of 95.72% and 97.43% for non-spiked saliva and spiked saliva classes, respectively (Fig. 1f). Also, tested on the training samples, the trained SVM classifier attained a F1-score of 100% for both non-spiked and spiked saliva classes. Low F1-score on the test samples and high F1-score on the training samples might be due to over-fitting of the classifier on the dataset. Therefore, to refute the over-fitting related confusion we also tested the performance of the classifier using the 5-fold cross-validation scheme. On 140 samples of non-spiked and spiked saliva classes, following the 5-fold cross-validation testing scheme the binary SVM classifier attained a mean F1-score of 92.85% and 100% for non-spiked and spiked saliva classes respectively. The confusion matrix of cross validation (Fig. S3e) reports that in the 5-fold experiment the binary SVM classifier correctly classified all the samples from the spiked saliva. However, out of 70 samples from the non-spiked saliva class, the classified misclassified five samples as spiked saliva class. The training and testing data details are given in S1.

Most of the biomolecules have characteristic Raman fingerprints. Since our approach could segregate the addition of a biomolecule like BSA in a lower concentration from complex bio fluid like saliva, we believe that any significant change in the saliva upon infection with the microbial species like bacteria or virus could also be differentiated effectively. The spiking study was extended with viral spike proteins. Coronavirus spike proteins from SARS-CoV-2, SARS-CoV, and MERS-CoV were employed, that represents the ACE-receptor binding domain leading to the viral infection and further replication. The Raman spectral pattern displayed a difference between the viral sub types and hence classified as three separate clusters up on PCA, LDA and SVM analysis (Fig. 2a, S3f, g). We observed that two components of kernel PCA were able to capture the differences in the Raman spectral data of SARS-CoV-2, SARS-CoV-1, and MERS-CoV spike proteins. Further, for the classification between these three spike proteins, a three class SVM classifier was trained using the features extracted by the kernel PCA for Raman spectral data. Evaluated on the test samples, the trained SVM classifier attained an F1-score of 100%, 100%, and 100% for SARS-COV-2, SARS-CoV, and MERS-CoV proteins, respectively (Fig. 2b). Also, evaluated on the training samples, the traine multi-class SVM also achieved an F1-score of 100% for each of the three protein classes. Apart from the train-test testing protocol, we also analyzed the performance of the 3-class SVM classifier using 5-fold cross-validation testing protocol. On the three-class protein dataset, following the 5-fold cross-validation scheme, the multi-class SVM classifier attained an F1-score of 100% for each of the three classes. The classification result of the classifier in terms of confusion matrix (Fig. S3h) shows that the classifier correctly classified all the samples from the three-class protein dataset. The training and testing data details are given in S2. Raman signatures at 540, 713, 760, 1002, and 1365 cm^{-1} indicate different assignments corresponding to amino acids cysteine, methionine, tryptophan, phenylalanine (Table S3). Variations were observed in the spectral pattern between the three viral proteins. Good classification and predictions made with the SVM classifier illustrated the capability of the current strategy to identify and differentiate viral infections by making use of unique Raman fingerprints. Recently, we have explored SERS modality to assess intermolecular interactions between suitably configured Tröger's base (TBs) and SARS-CoV-2spike proteins and ACE2 wherein the geometry of TBs was found to match the binding domain of SARS-CoV-2 and ACE2 [22].

To test whether our model could identify the presence of viral load roughly corresponding to 1-100 μg (10 [9]–10^{11} virions) in a clinically relevant concentration [23], a concentration-dependent study was done with saliva spiked with the S1 spike protein unit of SARS-CoV-2 (Fig. S4a, b). Our studies confirmed that even at a lower concentration (1 μg/mL) separation is possible. This observation clearly indicates the capacity of this framework to detect SARS-CoV-2 infection from the saliva even at an early stage of disease transmission [24]. As a next step, the study was per-formed with six healthy subjects spiked with SARS-CoV-2 (Fig. S4c, d).The spectral mapping and kernel PCA revealed that some extracted feature points of non-spiked protein were overlapped with the extracted feature points for spiked saliva. This may be due to heterogeneity of the saliva samples of the subjects considered for this study. Later, SVM classifier was trained using the 2 component feature extracted by kernel PCA. Evaluated on the test samples, the trained binary SVM classifier achieved an F1-score of 88.36% and 83.80% for saliva and spiked saliva classes, respectively (Fig. 2c, d). Besides, evaluated on the training samples, the trained classifier attained an F1-score of 89.23% and 100% for the saliva and spiked saliva classes, respectively. Also, on this dataset, the 5-fold cross-validation experiment achieved a mean F1-score of 93.33% and 100% for the saliva and spiked saliva classes, respectively. The confusion matrix result of the 5-fold cross-validation experiment (Fig. S4e) displays that the binary SVM classifier correctly classified all samples from the spiked saliva class. However, the classifier misclassified six samples from the saliva class into spiked saliva class. To further validate the findings, the same study was extended with a cohort size of 12 healthy subjects. Evaluated on the higher test and train samples in the dataset, the binary SVM classifier attained an F1-score of 100% for both saliva and spiked saliva classes (Fig. 2e, f). Besides, on this dataset, the 5-fold cross-validation experiment also attained an F1-score of 100% for both saliva and spiked saliva

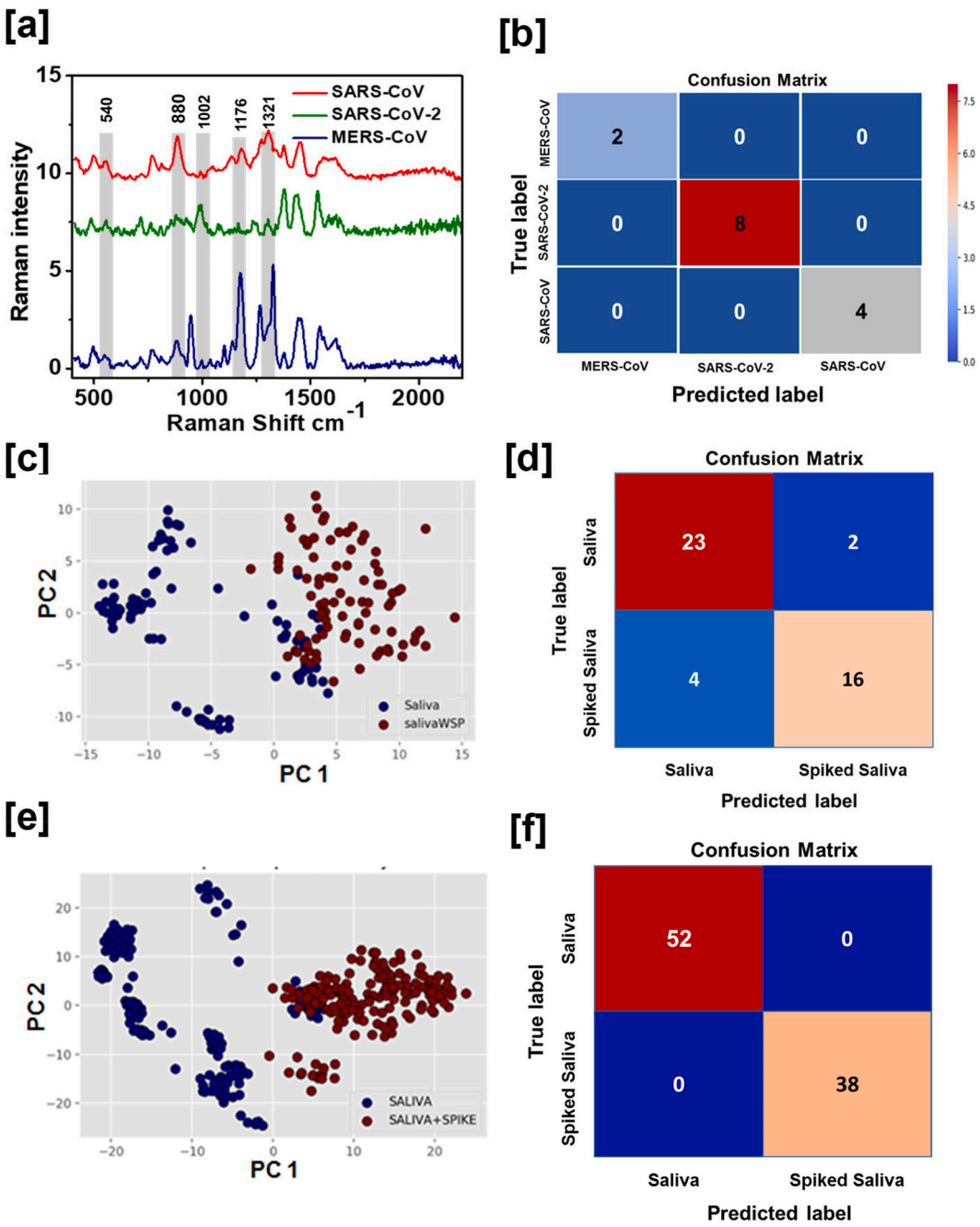


Fig. 2. [a] Staked Raman spectra and the corresponding [b] SVM analysis of three different Coronavirus spike proteins. [c] PCA and [d] SVM analysis of six subjects upon spiking of S1 protein of SARS-CoV-2. [e] PCA and [F] SVM analysis of twelve sub-jects upon spiking of S1 protein of SARS-CoV-2.

class. The classification result of the 5-fold cross-validation experiment in terms of confusion matrix (Fig. S4f) clearly illustrates that the binary SVM classifier correctly classified all samples from the saliva and spiked saliva classes. The training and testing data details are given in S3 and S4.

Unlike the other SARS virus diseases, the viral load of SARS-CoV-2 is highest during the first week after symptom onset and could be detected in the saliva as long as 20 days after the onset of symptoms. We next tested whether our approach could segregate the three corona virus spike proteins (SARS-CoV-2, SARS-CoV, and MERS-CoV) from the healthy saliva (Fig. S5a). We have extracted the prominent features from the Raman spectral data of saliva, and 3 different viral proteins using kernel PCA which proves that the plots have different characteristics. To classify the given samples in the above-mentioned 4-classes, an SVM classifier was trained using 2 components of extracted features from the kernel PCA. Evaluated on the test samples from the dataset, the trained SVM classifier achieved an F1-score of 79.51%, 100%, 85.71%, and 100% for saliva, MERS-CoV, SARS-CoV, and SARS-CoV-2 class, respectively (Fig. S5b, c). Later, studies were performed on nine healthy samples spiked with all three viral proteins. Although, the kernel PCA separates the most of the samples of four categories but in case of MERS-CoV and SARS-CoV spiked samples there is some overlap with the saliva samples which can be attributed to the less pronounced signature spectral features between these classes. Like-wise, the trained SVM classifier achieved F-score of 87.79% for saliva, 100% for SARS-CoV-2, 75.4% for SARS-CoV, and 85.71% for MERS-CoV (Fig. 3 a-c). Also, evaluated on the training samples, the classifier attained an F1-score of 97.10%, 98.14%, 100%, and 100% for saliva, saliva + MERS-CoV, saliva

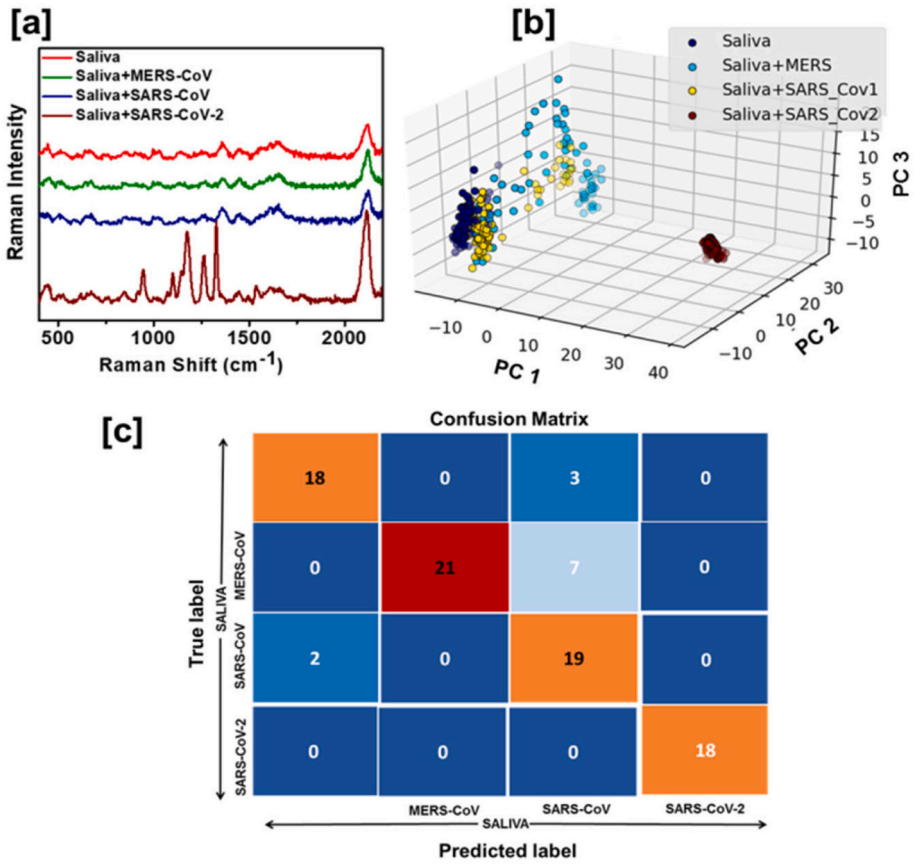


Fig. 3. Studies with three corona viral proteins spiked with saliva. [a] Stacked Raman spectra, [b] Kernel PCA and [c] Confusion matrix of 4-class SVM classifier.

+ SARS-CoV1, and saliva + SARS-CoV2 class, respectively. Besides, the 5-fold cross-validation experiment using the 4-class SVM classifier on the dataset attained an 94.44%, 78.04%, 55.57%, and 86.20% for saliva, saliva + MERS-CoV, saliva + SARS-CoV1, and saliva + SARS-CoV2 class, respectively. The confusion matrix results of the experiment (Fig. S5d) shows that the classifier correctly classified 85 samples from the saliva class, 64 samples from the saliva + MERS-CoV class, 50 samples from the saliva + SARS-CoV1 class, and all the 90 samples from the saliva + SARS-CoV2 class. Since the spectral acquisition and statistical analyses are based on the fingerprint changes in the saliva sample, the model studies can diagnose, distinguish and predict any viral infection. However, this claim can be proved only after studies on clinical saliva samples positive for the three viral infections. The training and testing data details are given in S5.

3.4. Studies with the Saliva Samples of COVID-19 Patients

The significant discrimination generated from the AI based algorithm of saliva and saliva-spiked corona virus spike proteins encourage us to extend the investigation on clinically confirmed for SARS-CoV-2 patient samples. Out of the 18 individuals selected for the study, nine were COVID positive with varying degrees of viral load. The samples were given pseudo-codes; the experiment and the data processing were performed as blind sample analysis to avoid any bias during the procedure. There was a marked variation in the spectral pattern of positive patients upon comparison with the healthy subjects (Fig. 4 a-d). The Raman peaks at 450, 1002, 1224, 1453, 1586, and 2126 cm^{-1} displayed variation between the samples. The principal characteristic peaks of thiocyanate stretching are around 450 and 2126 cm^{-1}. The presence of thiocyanate (SCN^-) is prevalent in minimal level in body secretions like saliva, serum, and urine and considered as a biomarker for evaluating human health. Smoking habit is an additional source of thiocyanate accumulation. [25] The Raman band at 2126 cm^{-1} from the C—N stretch of thiocyanate has been found higher in oral cancer subjects. The protective effects of thiocyanate against microbial peroxidases have already been studied [26,27]. In COVID positive samples, the peaks corresponding to thiocyanates i.e. 450 and 2126 cm^{-1} were increased when compared to the healthy counterparts. Additionally, C—N stretching and CH_3 vibrations in protein backbone, C—H stretching of glycoproteins, mostly generated from mucins, at 1453 cm^{-1} were found to increase (Table S4).The mucins may be increased in COVID-19 samples attributing to the immune activation induced by the virus [28]. Training and testing data details are given in S6. We approached to signal the SAR-Cov-2 spike protein spike in saliva samples to generate a proof-of-concept model for discriminating saliva and spiked saliva for the patient sample analysis even though the model and real samples showed no overlap with the Raman peaks.

The feature extracted by using the Kernel PCA method deciphered a good separation of SARS-CoV-2 infected samples with the negative controls (Fig. 4 e). Further, a binary SVM classifier was trained using 2 components of extracted kernel PCA features and the trained SVM classifier achieved an F1-score of 94.73% and 95.28% for healthy and COVID19 infected patients, respectively (Fig. 4f). Also, evaluated on the training samples, the trained binary classified attained an F1-score of 100% each for the healthy and COVID19 infected patients. A 100% F1-score on the training samples of the healthy and COVID19 infected patients might be due to over-fitting. Therefore, we also evaluated the performance of the SVM classifier using the 5-fold cross-validation testing protocol. Tested using a 5-fold cross-validation scheme the trained binary SVM classifier attained a mean F1-score of 98.89% and 100% for the healthy and COVID19 infected patient. Also, the confusion matrix result of the experiment (Fig. S5e) illustrates that the trained classifier correctly classified all the samples from the COVID-19 infected patient. However, the classifier wrongly classified one sample from the

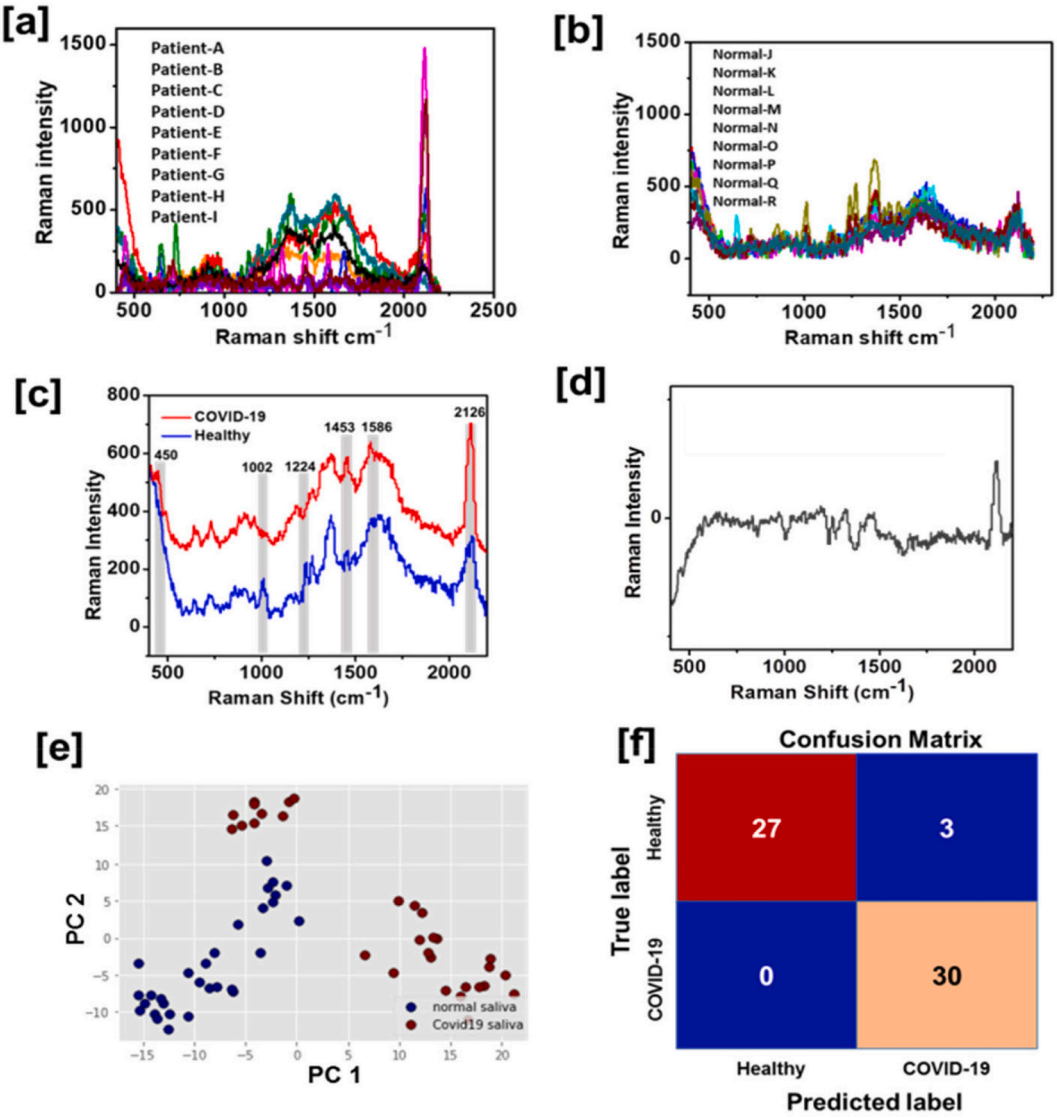


Fig. 4. Studies with the saliva samples of COVID patients. Raman spectra of nine [a] COVID-positive, [b] COVID-negative subjects, [c] representative stacked mean spectra and the [d] difference spectra between COVID-positive and healthy subjects. [e] PCA plot and [f] SVM analysis of the study.

healthy class into the COVID-19 infected class. Studies on SARS-CoV-2 illustrated the abundance of aromatic amino acids on the spike proteins are responsible for the interaction and subsequent internalization with ACE2. The difference in the Raman spectra could be due to the pathological features of the virus infection including the presence of circulating DNA, virus particles, shredded proteins, immune cells, and glycoproteins. Also, the presence of antibodies that are generated to counter the infection can modify the composition of saliva. The immune response will change the redox homeostasis and subsequently induce damage to the lipids on the plasma membrane. SARS-CoV-2 infection in the mouth could cause changes in saliva production or quality, contributing to symptoms of taste loss [29,30]. All the above-mentioned factors either alone or together may contribute to the altered salivary Raman spectral distribution pattern. The current investigation emphasizes on the feasibility of employing SERS as a straight forward, fast, non-invasive tool for the label-free discrimination of SARS-CoV-2 infection from saliva samples by employing colloidal AuNPs as SERS substrate. The specific Raman fingerprints of the biomolecules from the salivary SERS spectra, mainly viral proteins and nucleic acids are reflected which are subjected to classification models built-up by machine learning algorithms. Therefore, we could obtain a stand-alone classification from the Raman spectra of healthy vs COVID infected subjects which will ensure high accuracy, specificity, and sensitivity to the approach.

3.5. *Studies with the Saliva Samples of COVID-19 Recovered Subjects*

The clinical pathophysiology of SARS-CoV-2 infection encompasses asymptomatic infection, mild to severe fever, myalgias, fatigue and viral pneumonia. Many of these symptoms are persisting even after the patient has recovered from the infection [31]. Hence, follow-up studies to monitor the health status of people recovered from COVID-19 infection warrant at most urgency. Due to the advantages of our proposed system in terms of non-invasive sample collection procedure, and rapid disbursement of the results, we decided to check the label-free Raman fingerprints of a few saliva samples of COVID-19 recovered subjects. Raman spectra were acquired from the saliva samples after one month of their clinically recorded date of recovery from SARS-CoV-2 infection. Well resolved Raman spectra were subjected to PCA upon comparison with that of the healthy controls. The separation of both the groups as distinctly arranged clusters underlined the fact that bio-molecular changes in the saliva have not attained a healthy composition soon after recovery from the infection. However, the spectral patterns of the recovered subjects are more aligned towards healthy controls when

compared with that of positive control spectra. The follow-up investigation was made after two months of recovery. PCA illustrated that one subject aligned within the negative control cluster and another one is within the margin of negative control samples, while the other two subjects are clustered outside the negative control samples. The inclusion of two subjects with the healthy controls possibly indicated the normalcy attained in the biochemical composition of saliva over a period of time. The saliva samples of the same subjects were studied after 6 months of recovery wherein the recovered subjects and healthy controls aligned as a single cluster (Fig. 5a-c, S6, S7). Later, spectral analysis was performed with the saliva of another set of subjects after 15 days of recovery (Fig. S8a). The finding that both the classes were aligned as separate clusters confirmed the fact that the salivary composition of patients will be normalized only after a period of time. Additionally, it is interesting to notice that the thiocyanate peak at 450 and 2126 $cm^{-1}$is getting decreased for all the cases except one sample. The role of thiocyanate in viral infections must be further assessed with large number of samples to validate its potential as a valid biomarker for differentiating healthy and infected individuals. The representative stacked mean spectra and difference spectra between COVID-positive, COVID-recovered and healthy subjects are depicted in Fig. 5d and e. SVM classifier achieved a prediction accuracy of 95.38% and f1-score of

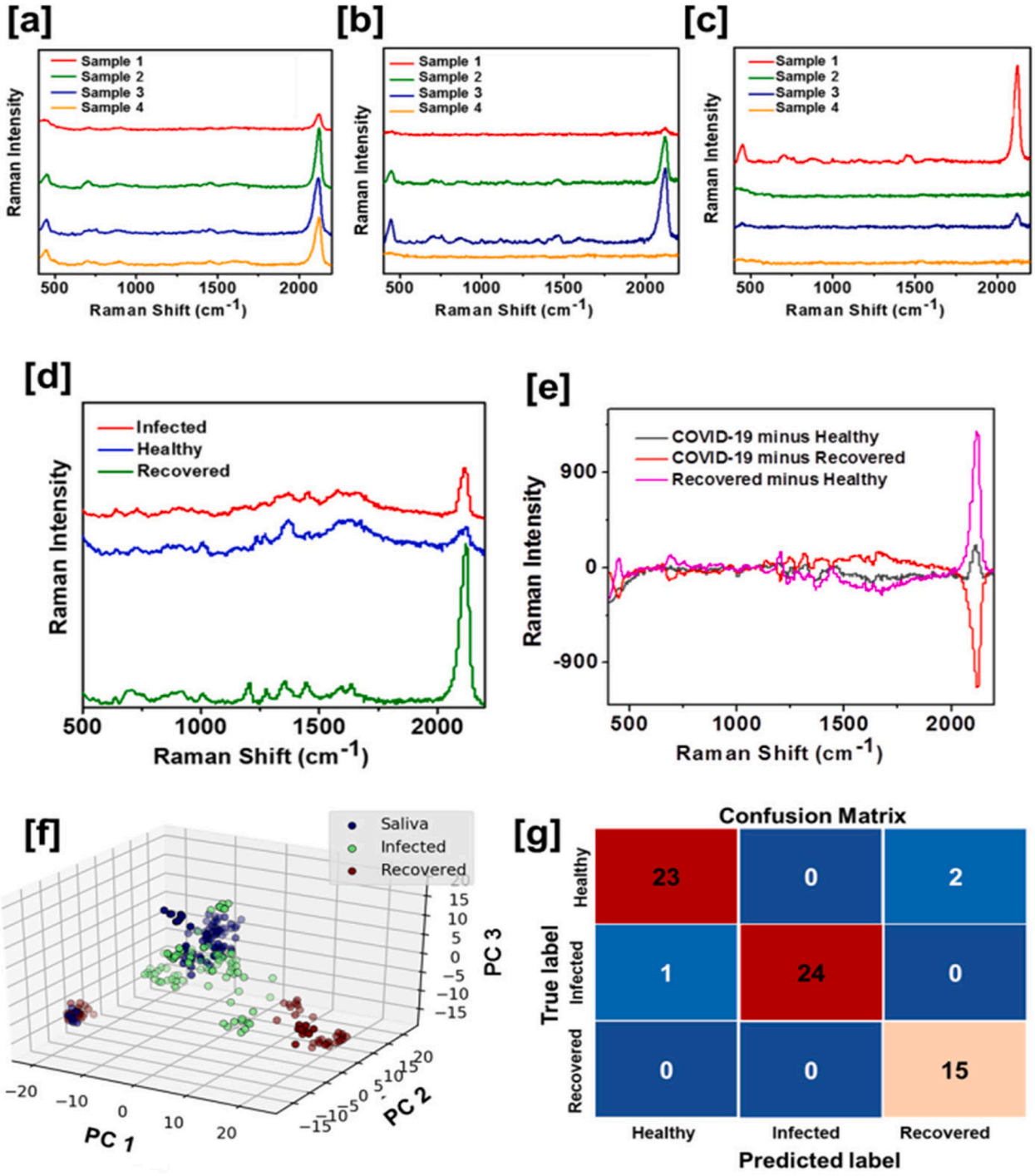


Fig. 5. Raman spectral analysis of COVID-recovered subjects (4 numbers) after [a] one month, [b] two months and [c] six months of recovery. [d] Representative stacked mean spectra and the [e] difference spectra between COVID-positive, COVID-recovered and healthy subjects. [f] PCA plot and [g] SVM analysis of the study.

94% for healthy subjects, f1-score of 98% for COVID-19 infected patients and f1-score of 94% for COVID-19 recovered patients (Fig. 5f, g). However, evaluated on the training samples of the dataset, the trained multi-class SVM classifier attained an F1-score of 98.66%, 100%, and 100% for the healthy, infected, and recovered samples from the patients. Also, the SVM classified trained using the 5-fold cross-validation scheme on the dataset attained a mean F1-score of 90%, 96.67%, and 100% for the healthy, infected, and recovered samples in the dataset. The classification result of the experiment in terms of confusion matrix (Fig. S8b) displays that the trained multi-class SVM classifier correctly classified all the samples from the recovered class. However, out of 100 samples from the healthy class the classifier correctly classified 90 samples. Also, the classifier correctly classified 87 samples out of 90 from the infecte class in the dataset. Training and testing data details are given in S7.

The characteristics of the subjects employed in this investigation in terms of sex, age, COVID infection status, and other important characteristics are illustrated in Tables S5-S7. The antibody response to SARS-CoV-2 infection has been broadly examined in the blood samples of COVID-19 recovered patients. Recently, the anti-SARS-CoV-2 antibody response over 115 days in the serum and saliva of COVID-19 recovered patients was investigated which revealed that saliva retains a good percentage of antibodies and hence can be used for monitoring the immune response to SARS-CoV-2 infection. Notably, along with a variety of factors, the presence of anti-SARS-CoV-2 anti-bodies could also be responsible for the discrimination of recovered patients from healthy controls [32,33]. Recently, a Raman-based classification model using saliva samples was able to discriminate COVID-19 patients with accuracy, precision, sensitivity and specificity [34]. Having admitted the fact of the very small sample size, the findings of this investigation are crucial in the backdrop of the constant reminders by the WHO for the need for follow-ups in the health status of recovered subjects. If our approximations hold true, patients need to take care of health with regular check-ups up to a stage where their salivary Raman spectral patterns return to a normal signature. Since sample collection and subsequent experimental procedures are simple to follow, a large cohort study on COVID19 recovered subjects is our immediate plan of action.

4. Conclusion

Clinical diagnostics for SARS-CoV-2 infection was accomplished using salivary label-free SERS fingerprinting with a hand-held Raman spectrometer and aided with artificial intelligence. The proof-of-concept was systematically optimized and initially demonstrated with spiking studies on receptor binding domain in viral proteins. Raman spectral data was further decoded, segregated, and effectively managed with the aid of machine learning algorithms. Later, salivary Raman fingerprint analysis on COVID-positive samples produced separation with high accuracy, specificity, and sensitivity form negative saliva. The current strategy also illustrated a database for different stages of patient recovery with varying Raman signatures. The higher prediction accuracy obtained could be due to the small sample size employed. Owing to the severe spreading nature of the diseases through the respiratory droplets and saliva, we were forced to limit the sample size. Moreover, the degree of infection and associated complexities were severe for the COVID-positive subjects and hence the viral load might also be very high. The accuracy rate may decrease if the sample size is significantly increased or include infected patients without symptoms and associated health problems. The currently employed methodology with a rapid and highly sensitive and non-invasive diagnostic procedure can assure patient comfort. Moreover, this technically proved methodology can be extended for large scale population screening forSARS-CoV-2 infection and surveillance study of COVID recovery process.

Author Contributions

VK, MMJ and SK undertook the synthesis of SERS substrate and all experimental studies and contributed in intellectual content. HS and SS contributed for Machine Learning /AI based Analysis of Raman data. RS, RM, TE, AA, SDP, YK and JJ contributed for analysis of Raman Data and preparation of training and testing dataset for developing Machine Learning/AI Algorithms. They also contributed for the preparation of substrate devices. RPS and SS. performed PCA and LDA analysis of the Raman spectral data using in-house software code developed by them. IY was involved in the planning and execution of the project, undertook patient sample handling and overall, in charge of ethical clearance and VA, BB, RD, TI, SDKL, NA and SMS are responsible of ethical clearance of the study. All authors discussed the results and commented on the manuscript. PSD was involved in designing and monitoring of experiments for saliva sample. KKM was responsible for the overall experiment planning, direction and coordination of the project.

Funding Sources

This work was financially supported by Council of Scientific and Industrial Research (CSIR: MLP0047), India.

Declaration of Competing Interest

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

Acknowledgment

This is a collaborative work between CSIR and the Research Unit, Department of Surgery, Government Medical College, Thiruvananthapuram, Kerala, India. Authors thank CSIR (MLP0047), Government of India for the research funding. K.K.M. thank the Director, CSIR-NIIST for the suggestions and progress evaluation of the work. Authors wish to thank all the participants who voluntarily provided samples for the study.

Appendix A. Supplementary data

The Supporting Information is available free of charge on the Synthetic details, experimental procedures, and supporting figures.

References

[1] D. Wrapp, N. Wang, K.S. Corbett, J.A. Goldsmith, C.-L. Hsieh, O. Abiona, B. S. Graham, J.S. McLellan, Cryo-EM structure of the 2019-NCoV spike in the Prefusion conformation, Science 367 (6483) (2020) 1260–1263, https://doi.org/10.1126/science.abb2507.

[2] J. Shang, G. Ye, K. Shi, Y. Wan, C. Luo, H. Aihara, Q. Geng, A. Auerbach, F. Li, Structural basis of receptor recognition by SARS-CoV-2, Nature 581 (7807) (2020) 221–224, https://doi.org/10.1038/s41586-020-2179-y.

[3] R. Wang, Y. Hozumi, G.-W.W. Changchuan Yin, Mutations on COVID-19 diagnostic targets, Genomics 112 (2020) 5204–5213.

[4] FDA, Policy for evaluating impact of viral mutations on COVID-19 tests, FDA No. February (2021) 1–13.

[5] N. Ravi, D.L. Cortade, E. Ng, S.X. Wang, Since January 2020 Elsevier Has Created a COVID-19 Resource Centre with Free Information in English and Mandarin on the Novel Coronavirus COVID- 19. The COVID-19 Resource Centre Is Hosted on Elsevier Connect, the Company's Public News and Information, 2020. No. January.

[6] R. Xu, B. Cui, X. Duan, P. Zhang, X. Zhou, Q. Yuan, Saliva: potential diagnostic value and transmission of 2019-NCoV, Int. J. Oral Sci. 12 (1) (2020), https://doi.org/10.1038/s41368-020-0080-z.

[7] D. Ter-Ovanesyan, T. Gilboa, R. Lazarovits, A. Rosenthal, X. Yu, J.Z. Li, G. M. Church, D.R. Walt, Ultrasensitive measurement of both SARS-CoV-2 RNA and antibodies from saliva, Anal. Chem. 93 (13) (2021) 5365–5370, https://doi.org/10.1021/acs.analchem.1c00515.

[8] M.M. Joseph, N. Narayanan, J.B. Nair, V. Karunakaran, A.N. Ramya, P.T. Sujai, G. Saranya, J.S. Arya, V.M. Vijayan, K.K. Maiti, Exploring the margins of SERS in practical domain: an emerging diagnostic modality for modern biomedical applications, Biomaterials 181 (2018) 140–181, https://doi.org/10.1016/j.biomaterials.2018.07.045.

[9] T.D. Payne, S.J. Klawa, T. Jian, S.H. Kim, M.J. Papanikolas, R. Freeman, Z. D. Schultz, Catching COVID: engineering peptide-modified surface-enhanced

Raman spectroscopy sensors for SARS-CoV-2, ACS Sens. 6 (9) (2021) 3436–3444, https://doi.org/10.1021/acssensors.1c01344.
[10] H. Chen, S.G. Park, N. Choi, H.J. Kwon, T. Kang, M.K. Lee, J. Choo, Sensitive detection of SARS-CoV-2 using a SERS-based Aptasensor, ACS Sens. 6 (6) (2021) 2378–2385, https://doi.org/10.1021/acssensors.1c00596.
[11] M. Zhang, X. Li, J. Pan, Y. Zhang, L. Zhang, C. Wang, X. Yan, X. Liu, G. Lu, Ultrasensitive detection of SARS-CoV-2 spike protein in untreated saliva using SERS-based biosensor, Biosens. Bioelectron. 190 (2021), 113421, https://doi.org/10.1016/j.bios.2021.113421.
[12] S. Weng, H. Yuan, X. Zhang, P. Li, L. Zheng, J. Zhao, L. Huang, Deep learning networks for the recognition and quantitation of surface-enhanced Raman spectroscopy, Analyst 145 (14) (2020) 4827–4835, https://doi.org/10.1039/d0an00492h.
[13] V. Karunakaran, V.N. Saritha, M.M. Joseph, J.B. Nair, G. Saranya, K.G. Raghu, K. Sujathan, K.S. Kumar, K.K. Maiti, Diagnostic Spectro-cytology revealing differential recognition of cervical cancer lesions by label-free surface enhanced Raman fingerprints and chemometrics, Nanomedicine 29 (2020) 102276.
[14] S.-J. Baek, A. Park, Y.-J. Ahn, J. Choo, Baseline correction using asymmetrically reweighted penalized least squares smoothing, Analyst 140 (1) (2015) 250–257, https://doi.org/10.1039/c4an01061b.
[15] C.C. Chang, C.J. Lin, LIBSVM: a library for support vector machines, ACM Trans. Intell. Syst. Technol. 2 (3) (2011), https://doi.org/10.1145/1961189.1961199.
[16] G. Yang, J. Nanda, B. Wang, G. Chen, D.T. Hallinan, Self-assembly of large gold nanoparticles for surface-enhanced Raman spectroscopy, ACS Appl. Mater. Interfaces 9 (15) (2017) 13457–13470, https://doi.org/10.1021/acsami.7b01121.
[17] L. Jiang, T. You, P. Yin, Y. Shang, D. Zhang, L. Guo, S. Yang, Surface-enhanced Raman scattering spectra of adsorbates on Cu2O nanospheres: charge-transfer and electromagnetic enhancement, Nanoscale 5 (7) (2013) 2784–2789, https://doi.org/10.1039/c3nr33502j.
[18] P.T. Sujai, M.M. Joseph, G. Saranya, J.B. Nair, V.P. Murali, K.K. Maiti, Surface charge modulates the internalization: vs. penetration of gold nanoparticles: comprehensive scrutiny on monolayer cancer cells, multicellular spheroids and solid tumors by SERS modality, Nanoscale 12 (13) (2020) 6971–6975, https://doi.org/10.1039/d0nr00809e.
[19] G. Calado, I. Behl, A. Daniel, H.J. Byrne, F.M. Lyng, Raman spectroscopic analysis of saliva for the diagnosis of oral cancer: a systematic review, Transl. Biophoton. 1 (1–2) (2019) 1–10, https://doi.org/10.1002/tbio.201900001.
[20] Z. Movasaghi, S. Rehman, I. Rehman, Raman spectroscopy of biological tissues, Appl. Spectrosc. Rev. 42 (5) (2007) 493–541, https://doi.org/10.1080/05704920701551530.
[21] A. Hole, G. Tyagi, A. Deshmukh, R. Deshpande, V. Gota, P. Chaturvedi, C. M. Krishna, Salivary Raman spectroscopy: standardization of sampling protocols and stratification of healthy and oral cancer subjects, Appl. Spectrosc. 75 (5) (2021) 581–588, https://doi.org/10.1177/0003702820973260.
[22] A. Linet, M.M. Joseph, M. Haritha, K. Shamna, S. Varughese, P.S. Devi, C. H. Suresh, K.K. Maiti, I. Neogi, De Novo design and synthesis of boomerang-shaped molecules and their in silico and SERS-based interactions with SARS-CoV-2 spike protein and ACE2, New J. Chem. 45 (38) (2021) 17777–17781, https://doi.org/10.1039/d1nj02955j.
[23] R. Sender, Y.M. Bar-On, S. Gleizer, B. Bernshtein, A. Flamholz, R. Phillips, R. Milo, The total number and mass of SARS-CoV-2 virions, Proc. Natl. Acad. Sci. U. S. A. 118 (25) (2021) 1–9, https://doi.org/10.1073/pnas.2024815118.
[24] C.J. Seneviratne, P. Balan, K.K.K. Ko, N.S. Udawatte, D. Lai, D.H.L. Ng, I. Venkatachalam, K.S. Lim, M.L. Ling, L. Oon, B.T. Goh, X.Y.J. Sim, Efficacy of commercial mouth-rinses on SARS-CoV-2 viral load in saliva: randomized control trial in Singapore, Infection 49 (2) (2021) 305–311, https://doi.org/10.1007/s15010-020-01563-9.
[25] A. Spagnolo, S. Torsello, G. Morisi, E. Petrozzi, R. Antonini, G. Ricci, G.C. Urbinati, A. Menotti, Serum thiocyanate levels as an objective measure of smoking habits in epidemiological studies, Eur. J. Epidemiol. 4 (2) (1988) 206–211, https://doi.org/10.1007/BF00144753.
[26] A. Fălămaș, H. Rotaru, M. Hedeșiu, Surface-enhanced Raman spectroscopy (SERS) investigations of saliva for oral cancer diagnosis, Lasers Med. Sci. 35 (6) (2020) 1393–1401, https://doi.org/10.1007/s10103-020-02988-2.
[27] J. Tenovuo, H. Larjava, The protective effect of peroxidase and thiocyanate against hydrogen peroxide toxicity assessed by the uptake of [3H]-thymidine by human gingival fibroblasts cultured in vitro, Arch. Oral Biol. 29 (6) (1984) 445–451.
[28] M. Bose, B. Mitra, P. Mukherjee, Mucin signature as a potential tool to predict susceptibility to COVID-19, Phys. Rep. 9 (1) (2021) 1–10, https://doi.org/10.14814/phy2.14701.
[29] G.A. Poland, I.G. Ovsyannikova, R.B. Kennedy, SARS-CoV-2 immunity: review and applications to phase 3 vaccine candidates, Lancet 396 (10262) (2020) 1595–1606, https://doi.org/10.1016/S0140-6736(20)32137-1.
[30] Y. Okada, K. Yoshimura, S. Toya, M. Tsuchimochi, Pathogenesis of taste impairment and salivary dysfunction in COVID-19 patients, Jpn. Dent. Sci. Rev. 57 (2021) 111–122, https://doi.org/10.1016/j.jdsr.2021.07.001.
[31] V. Higgins, D. Sohaei, E.P. Diamandis, I. Prassas, COVID-19: from an acute to chronic disease? Potential long-term health consequences, Crit. Rev. Clin. Lab. Sci. 58 (5) (2021) 297–310, https://doi.org/10.1080/10408363.2020.1860895.
[32] E. Kowitdamrong, T. Puthanakit, W. Jantarabenjakul, E. Prompetchara, P. Suchartlikitwong, O. Putcharoen, N. Hirankarn, Antibody responses to SARS-CoV-2 in patients with differing severities of coronavirus disease 2019, PLoS One 15 (2020) 1–11, https://doi.org/10.1371/journal.pone.0240502.
[33] Y. Chen, L. Li, SARS-CoV-2: virus dynamics and host response, Lancet Infect. Dis. 20 (5) (2020) 515–516, https://doi.org/10.1016/S1473-3099(20)30235-8.
[34] C. Carlomagno, D. Bertazioli, A. Gualerzi, S. Picciolini, P.I. Banfi, A. Lax, E. Messina, J. Navarro, L. Bianchi, A. Caronni, F. Marenco, S. Monteleone, C. Arienti, M. Bedoni, COVID-19 salivary Raman fingerprint: innovative approach for the detection of current and past SARS-CoV-2 infections, Sci. Rep. 11 (1) (2021) 1–13.