Received: 30 April 2023 | Revised: 11 September 2023 | Accepted: 22 September 2023

DOI: 10.1002/biot.202300191

Biotechnology Journal

RESEARCH ARTICLE

Characterization and discrimination of spike protein in SARS-CoV-2 virus-like particles via surface-enhanced Raman spectroscopy

Munevver Akdeniz[1,2] | Zakarya Al-Shaebi[1,2] | Mine Altunbek[3] | Canan Bayraktar[4] | Alisan Kayabolen[4,5] | Tugba Bagci-Onder[4] | Omer Aydin[1,2,6,7]

[1]Department of Biomedical Engineering, Erciyes University, Kayseri, Turkey

[2]Nanothera Lab, Drug Application and Research Center (ERFARMA), Erciyes University, Kayseri, Turkey

[3]Department of Chemical Engineering, University of Massachusetts, Lowell, Massachusetts, USA

[4]Koç University Research Center for Translational Medicine (KUTTAM), Koç University, Istanbul, Turkey

[5]McGovern Institute for Brain Research at MIT, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA

[6]Clinical Engineering Research and Implementation Center (ERKAM), Erciyes University, Kayseri, Turkey

[7]Nanotechnology Research and Application Center (ERNAM), Erciyes University, Kayseri, Turkey

Correspondence
Omer Aydin, Department of Biomedical Engineering, Erciyes University, 38039, Kayseri, Turkey.
Email: biomer@umich.edu, omeraydin@erciyes.edu.tr

Funding information
Bilimsel Araştırma Projeleri, Erciyes Üniversitesi, Grant/Award Number: FYL-2021-10780

Abstract

Non-infectious virus-like particles (VLPs) are excellent structures for development of many biomedical applications such as drug delivery systems, vaccine production platforms, and detection techniques for infectious diseases including SARS-CoV-2 VLPs. The characterization of biochemical and biophysical properties of purified VLPs is crucial for development of detection methods and therapeutics. The presence of spike (S) protein in their structure is especially important since S protein induces immunological response. In this study, development of a rapid, low-cost, and easy-to-use technique for both characterization and detection of S protein in the two VLPs, which are SARS-CoV-2 VLPs and HIV-based VLPs was achieved using surface-enhanced Raman spectroscopy (SERS). To analyze and classify datasets of SERS spectra obtained from the VLP groups, machine learning classification techniques including support vector machine (SVM), k-nearest neighbors (kNN), and random forest (RF) were utilized. Among them, the SVM classification algorithm demonstrated the best classification performance for SARS-CoV-2 VLPs and HIV-based VLPs groups with 87.5% and 92.5% accuracy, respectively. This study could be valuable for the rapid characterization of VLPs for the development of novel therapeutics or detection of structural proteins of viruses leading to a variety of infectious diseases.

KEYWORDS
COVID-19, HIV, machine Learning, SARS-CoV-2, surface-enhanced Raman spectroscopy, virus-like particle

1 | INTRODUCTION

The COVID-19 outbreak, caused by the SARS-CoV-2 virus, posed a global challenge. SARS-CoV-2 has a single-stranded RNA and four key structural proteins: spike (S), nucleocapsid (N), membrane (M), and envelope (E).[1,2] The N protein encapsulates genetic material, while M and E proteins create the viral envelope. The S protein binds to host cells via ACE2 receptors, and its abundant presence is crucial for diagnosis, vaccines, and drugs.[3] Despite efforts to develop rapid detection and treatment methods, the virus's high contagion and pathogenicity have hindered progress.

VLPs for SARS-CoV-2 are promising for expanding virus studies with a lower biosafety level than required.[4] VLPs comprise of viral structural proteins. They are non-infectious replicates since molecules and

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.
© 2023 The Authors. Biotechnology Journal published by Wiley-VCH GmbH

do not contain all viral genetic materials, which make them promising for many biological applications such as drug, nucleic acid or protein delivery systems, and vaccine production.[5,6] Especially, they are preferred in vaccine development due to their immunogenicity.[7,8] VLPs can be produced in various cultures, including bacteria, mammalian cell lines, insects, yeast, and plant cells.[9] The characterization of biochemical composition after purification of VLP gives critical information such as presence of delivery material or representation of structural proteins, especially the correct display of immunogenic epitopes which are important for their functionality in therapeutic applications. VLPs currently are characterized by different methods such as transmission electron microscopy (TEM), nanoparticle tracking analysis (NTA), SDS-PAGE, Western blotting, enzyme-linked immunosorbent assay (ELISA),[10] or serological tests.[11] However, these methods have many drawbacks such as the need for complex equipment, requiring high-cost reagents (i.e., enzymes, antibodies, primers, buffers, polymerases), long processing times, or need for high amount of samples.[12] As a result of these flaws, present techniques are not neither appropriate nor sufficient for rapid characterization of VLPs.

Raman spectroscopy (RS) is a widely used vibrational spectroscopic technique. The principal of RS is based on inelastic scattering of laser light which reveals information about vibrational modes for a molecule.[13,14] However, the Raman scattering from biological molecules is weak that can cause loss of precious information in noise signals. The enhancement in Raman scattering can be achieved when the plasmonic particles in form of colloidal nanoparticles or nanostructured surfaces, are placed close proximity to analyte and this technique is called as SERS.[15,16] This enhancement is due to formation of hotspots around the localized surface plasmons of metallic substrates.[17,18] Colloidal nanoparticles such as silver (AgNPs) and gold (AuNPs) are commonly used SERS substrates, but use of nanoparticles can cause spectral fluctuations due to their uncontrolled and random aggregation.[19] Recently, a practical way to form reproducible hotspots is proposed by controlled aggregation of colloidal nanoparticles in presence of using inorganic salts such as NaCl, $MgSO_4$, and $CuSO_4$.[20,21] This method is relatively cost-effective and has a simple process for preparation of SERS substrate. For example, Lim et al.[22] compared different salts including $MgCl_2$, $CaCl_2$, and $CuSO_4$ to control aggregation of AuNPs enabled to reach more precious molecular information and detection of different strains of influenza viruses. In our previous study, we also reported SERS spectra of SARS-CoV-2 structural proteins with high sensitivity and reproducibility by the controlled aggregation of AuNPs in the presence of $CuSO_4$.[23] In another study, Bell et al. used $MgSO_4$ for the aggregation of AgNPs, which decreased the limit of detection for DNA/RNA mononucleotides under 1 ppm (1 μg mL^{-1}).[24]

There are limited studies for the characterization and detection of VLPs using SERS in the literature. Sitjar et al. reported the detection of biochemical structures in viral structures with a simple, cost-efficient sample preparation method.[25] They developed a SERS active substrate composed of gold nanoholes (AuNC) and Au nanoparticles on ZrO_2 nanobowls (AuNPs/pZrO_2) and demonstrated the detection of SARS-CoV-2 pseudovirus, vesicular stomatitis virus G (VSV-G), and pseudotype lentivirus. In another study, Lebedev et al. synthesized self-assembly plasmonic nanoclusters using VLPs and AuNP and demonstrated that plasmonic structures can be used in DNA detection in 0.25 ng μL^{-1} concentration.[26]

The identification of a simple molecule or discrimination of structural proteins among the multiple peaks in a SERS spectrum is quite challenging since a small change in a spectrum might indicate the emerging mutations.[27] Machine learning approaches are powerful tools to classify different groups by analyzing large SERS data sets by finding minimum variation within the groups while exploring maximum variation between the groups.[28–31]

In this study, we prepared two VLP groups: HIV-based VLPs produced with HIV-gag protein and SARS-CoV-2 VLPs produced from M and N proteins of SARS-CoV-2. To analyze and classify the two noninfectious VLP groups was used SERS, which is a rapid, cost-effective technique, AuNPs with 50 nm diameter size were used as SERS substrate and $CuSO_4$ was used to generate hot spots. SERS spectra were collected from four different types of SARS-CoV-2 VLPs and four different types of HIV-based VLPs. Multivariate analyses with traditional classifiers were applied to SERS spectra to extract precious information and effectively classify the groups. To the best of our knowledge, this is the first study that has been reported for characterization and discrimination of SARS-CoV-2 VLPs and HIV-based VLPs with SERS in the literature. This study can be a guide to better understand and characterize structure of the alternative viral structures to wild type virus.

2 | EXPERIMENTAL

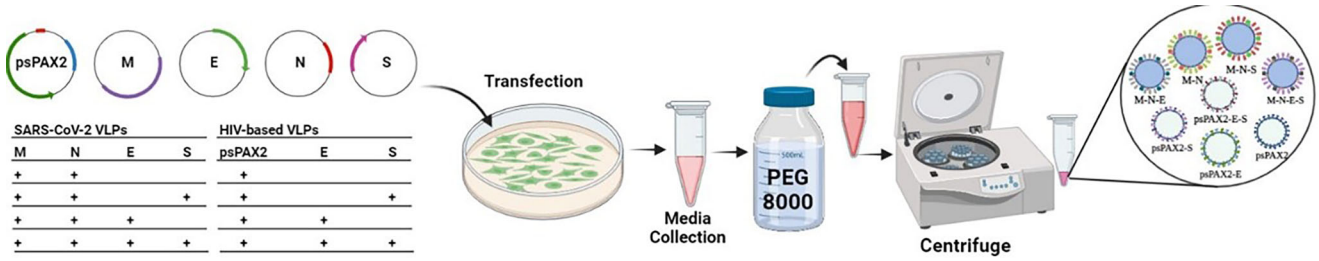
2.1 | Materials

Gold (III)Chloride hydrate ($HAuCl_4$), copper sulfate ($CuSO_4$), sodium citrate and Poly(ethylene glycol), MW 8000 were purchased from Merck (Germany), and phosphate buffer saline (PBS) was purchased from Gibco. Poly(ethyleneimine), linear, MW 25000 was purchased from Polysciences.

2.2 | Virus-like particles (VLPs) preparation and characterization

HEK293T cells were purchased from American Type Culture Collection (ATCC, USA), and have been tested regularly for mycoplasma contamination. They were cultured in DMEM with 10% FBS, and 1% Pen/Strep (all from Gibco, USA). One day before transfection, 6×10^6 of HEK293T cells were seeded on 10-cm plates. Then, they were transfected with appropriate plasmids using polyethyleneimine (PEI) transfection agent. Ten micrograms of total DNA and 40 μL of PEI were used for each 10-cm plates. Plasmid mixtures were prepared according to four different SARS-CoV-2 VLP groups: M-N which contains membrane and nucleocapsid protein of SARS-CoV-2 in structure, M-N-S

18607314, 2024, 1, Downloaded from https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/biot.202300191 by Koc University, Wiley Online Library on [15/02/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License



SCHEME 1 Schematic illustration of production of SARS-CoV-2 VLPs and HIV-based VLPs using PEI transfection agent and plasmids.

which contains membrane, nucleocapsid, and spike protein in structure, M-N-E which contain membrane, nucleocapsid, and envelope protein in structure, and M-N-E-S which contain all of the structural proteins of SARS-CoV-2, and according to four different HIV-based VLP groups: psPAX2, which does not contain the structural protein of SARS-CoV-2 in its structure and is used as a control group, psPAX2-S, which contains spike protein in its structure, psPAX2-E, which contains envelope proteins in its structure, and psPAX2-E-S, which contains both spike and envelope protein in its structure. After 14–16 h of transfection, media was changed with fresh media (DMEM with 10% FBS and 1% Pen/Strep). After 24 and 48 h of transfection, media collected and filtered through 0.45 μm syringe filters. They were concentrated by overnight incubation with polyethyleneglycol (PEG 8000, 50% w/v) at 4°C with a 10% final PEG ratio. VLPs were centrifuged for 20 min at 2500 rpm. Pellets were resuspended with cold PBS and stored at −80°C (Scheme 1).

To characterize VLP samples, total protein concentrations of concentrated VLP samples were determined via Pierce's BCA protein assay kit (Thermo Fisher Scientific, USA). Samples were mixed with 4X loading dye which is prepared by mixing 4 × Laemmli Sample Buffer (Bio-Rad, USA) with 2-mercaptoethanol in 9:1 ratio, and boiled at 95°C for 10 min. Equal protein amount of samples were loaded into gradient (4%–15%) SDS polyacrylamide gels (Mini-PROTEAN TGX Precast Gels, Bio-Rad, USA), and run at 30 mA for 70 min. Then, proteins were transferred to PVDF membrane by wet transfer at 4°C for 2 h. After transfer, membrane was stained with Ponceau to visualize total protein amount loaded from each sample and blocked with TBS-T (0.1% Tween-20) containing 5% non-fat dry milk for 1 h with gentle shaking at room temperature. Then, blocking buffer was replaced with primary anti-Spike antibody (99423, Cell Signaling Technology), diluted in TBS-T with 5% non-fat dry milk and incubated overnight by gently shaking at 4°C. Next day, antibody solution was removed, and membrane was washed three times with PBS-T for 5, 10, and 15 min, respectively. Then, membrane was incubated with secondary antibody diluted in TBS-T containing 5% non-fat dry milk for 1 h at RT and washed 3 times with PBS-T for 5, 10, and 15 min, respectively. Membrane was incubated with SuperSignal™ West Femto Maximum Sensitivity Substrate (Thermo Fisher Scientific, USA) for 5 min at dark, and visualized by Odyssey Fc Imaging System (LI-COR Biosciences, USA).

Concentrations of SARS-CoV-2 VLPs and HIV-based VLPs were determined using a nanoparticle tracking analyzer instrument (NTA-Malvern, Nanosight NS300, UK) equipped with a blue laser (405 nm) and NTA 3.4 software. Each sample was prepared by dilution 100-fold in PBS (viscosity of PBS = 1072). The camera level was adjusted manually, 60 s videos were captured at room temperature.

2.3 | Preparation of SERS substrates

Colloidal AuNPs with average 50 nm size were synthesized using the Turkevich method.[32] In brief, 5 mg of $HAuCl_4$ was dissolved in 50 mL ultrapure dH_2O and heated until boiling. A 300 μL sodium citrate (1%) solution was added rapidly into the boiling solution and the solution was kept boiling for 15 min. Absorption spectrum of the AuNPs was obtained using UV-Visible Spectrometer (Perkin Elmer, Lambda 25, USA) in the range of 300–700 nm. Zeta potential and hydrodynamic diameter of the AuNPs were determined using Zeta-sizer (Malvern, Nano ZS, UK). Scanning-Transmission Electron Microscope (STEM) images were obtained with ZEISS GEMINI 500 instrument (Germany).

SERS substrates were prepared by using our previously published protocol and the following literature.[22,23,33] Briefly, the synthesized AuNP colloidal solution was mixed with 10 mM $CuSO_4$ solution (50:1, v/v). The mixture was dropped onto CaF_2 slides and air-dried at room temperature before SERS measurements.

2.4 | SERS measurements

SARS-CoV-2 VLPs and HIV-based VLPs with a 4×10^7 PFU mL^{-1} concentration were dropped on SERS substrate and dried at room temperature for 5 min. SERS measurements were acquired WITech alpha M+ Raman Microscopy System (Germany). Samples were excited with a near-infrared 785 nm diode laser with a laser power of 30 mW under 50 × objective lens. Integration time was 2 s in all experiments. Each spectrum consists of 818 points in the range of 400–1800 cm^{-1}. Three independent measurements were carried out under the same experiment conditions.

2.5 | Data preprocessing and multivariate analysis

A total of 70 spectra were obtained from each group in the range of 400–1800 cm^{-1}. The training dataset used in this study consisted of a total of 60 spectra from each group. These spectra were randomly selected and used for training the model. Additionally, a separate test

18607314, 2024, 1, Downloaded from https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/biot.202300191 by Koc University, Wiley Online Library on [15/02/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

dataset was selected randomly, containing 10 spectra from each group, to evaluate the performance of the model after training. The random selection of spectra ensures that the training dataset is representative and avoids any bias in the model's learning process. The obtained SERS spectra were processed for noise reduction as fallow. Cosmic ray removal, background subtraction and Savitzky Golay smoothing. All spectral vectors were then normalized between 0 and 1. Finally, machine learning analysis was performed on Google Colab using the machine learning binary classifiers from the Scikit-learn python library. The kNN, RF, and SVM, including linear, Gaussian, and poly classification algorithms were used to classify SARS-CoV2 VLP groups and HIV-based VLP groups.

3 | RESULTS AND DISCUSSIONS

3.1 | SERS substrate characterization

The preparation of SERS substrate is critical step for acquisition of highly sensitive and reproducible SERS spectra. We followed a simple SERS substrate fabrication process. First, we synthesized AuNPs with an average 50 nm size since NPs around 50 nm were found better for enhancement of the SERS spectrum.[34,35] The size of AuNPs was around 50 nm characterized with the maximum absorption at 534 nm and hydrodynamic size distribution analysis (Figure S1A-B). The zeta potential of AuNPs was found −20 mV indicating that the nanoparticles are stable in colloidal suspension (Figure S1C).[36] $CuSO_4$ induced AuNPs aggregation by interacting with weakly bound citrate ions on nanoparticles as seen in Figure S1D.[24] After simply dropping AuNPs-$CuSO_4$ mixture on CaF_2, the SERS substrates became ready for the use in SERS measurements.

3.2 | Characterization of VLP structures

The characterization of the produced VLPs as in Scheme 1 was performed by western blotting. To confirm presence of the Spike as an envelope protein on VLPs, antibody against SARS-CoV-2 Spike protein was used for the blotting. Western images clearly show presence of the Spike protein on the VLPs packaged with Spike plasmid, namely psPAX2-Spike, psPAX2-E-Spike, M-N-Spike, M-N-E-Spike, while no bands have been observed in the S-VLPs which are packaged without envelopes or with only E envelope protein, namely psPAX2, psPAX2-E, M-N, M-N-E (Figure 1). As a loading control, Ponceau staining shows that similar amounts of total protein were loaded from each sample (Figure S2). These results are also consistent with measurements taken with the SERS spectra in the presence of spike protein (Figures 2 and 3).

3.3 | SERS spectra of SARS-CoV-2 VLPs and HIV-based VLPs

Four different groups of SARS-CoV-2 VLPs and HIV-based VLPs were simply dropped on the prepared SERS substrates in equal concentration (Figure S3) calculated following Equation (1):

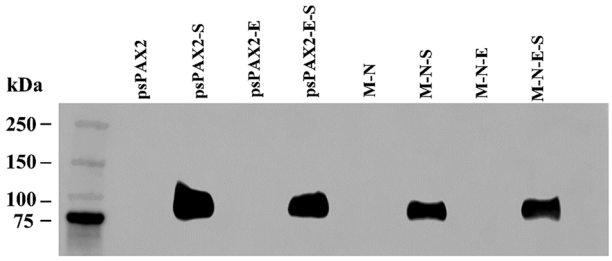


FIGURE 1 Characterization of SARS-CoV-2 VLPs and HIV-based VLPs with western blotting, the bands indicate spike protein presence in VLPs structures.

$$dropping\ volume = \frac{required\ concentration \times 1000\ \mu L}{Stock\ concentration} \quad (1)$$

where required concentration represents desired concentration (4×10^7 PFU mL^{-1}), stock concentration is the measured concentration of each sample. After calculation of concentration were mixed with PBS to drop equal volume, a total of 70 spectra were collected from each group to decipher molecular information. The mean normalized SERS spectra of each group are compared in Figure 2 for SARS-CoV-2 VLPs and Figure 3 for HIV-based VLPs. The blank substrate did not display any background peaks demonstrating that no impurities contributed to the SERS spectra of SARS-CoV-2 VLPs or HIV-based VLPs (Figure S4). Reproducibility of spectra is demonstrated by spot-to-spot and batch-to-batch measurements given in Figure S5–S6 for SARS-CoV-2 VLPs and in Figure S7–S8 for HIV-based VLPs. SARS-CoV-2 VLPs and HIV-based VLPs showed reproducible SERS spectra. The percentage relative standard deviation (RSD %) was determined to show the reproducibility of the SERS spectra of SARS-CoV-2 VLPs and HIV-based VLPs, and calculated according to the following Equation (2):

$$RSD\ \% = \frac{Standart\ Deviation}{Mean} \times 100 \quad (2)$$

The RSD coefficient was calculated according to the two peaks at 842 and 858 cm^{-1} related to discrimination of S protein in all groups. Spectra of reproducibility were collected under the same experimental conditions. RSD coefficients for spot-to-spot and batch-to-batch measurements are given in Table S1 for SARS-CoV-2 VLPs and in Table S2 for HIV-based VLPs. Thus, our results show reliable reproducibility for obtained spectra.

The SERS spectra obtained from SARS-CoV-2 VLPs are depicted in Figure 2A. The peak assignments that contributed to the structural changes are shown in Table S3. The peaks found on the spectra of SARS-CoV-2 VLPs and HIV-based VLPs are similar to the peaks in the spectra collected from both inactivated versions of human coronavirus and S, E proteins by the study of the Sanchez et al.[37] The prominent peaks arise at 535, 581, 842, 858, 935, 1063, 1125, 1142, 1233, 1278, 1395, 1449, and 1481 cm^{-1}, which exhibited different patterns between four different SARS-CoV-2 VLPs including M-N, M-N-S, M-N-E, M-N-E-S.

18607314, 2024, 1, Downloaded from https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/biot.202300191 by Koc University, Wiley Online Library on [15/02/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

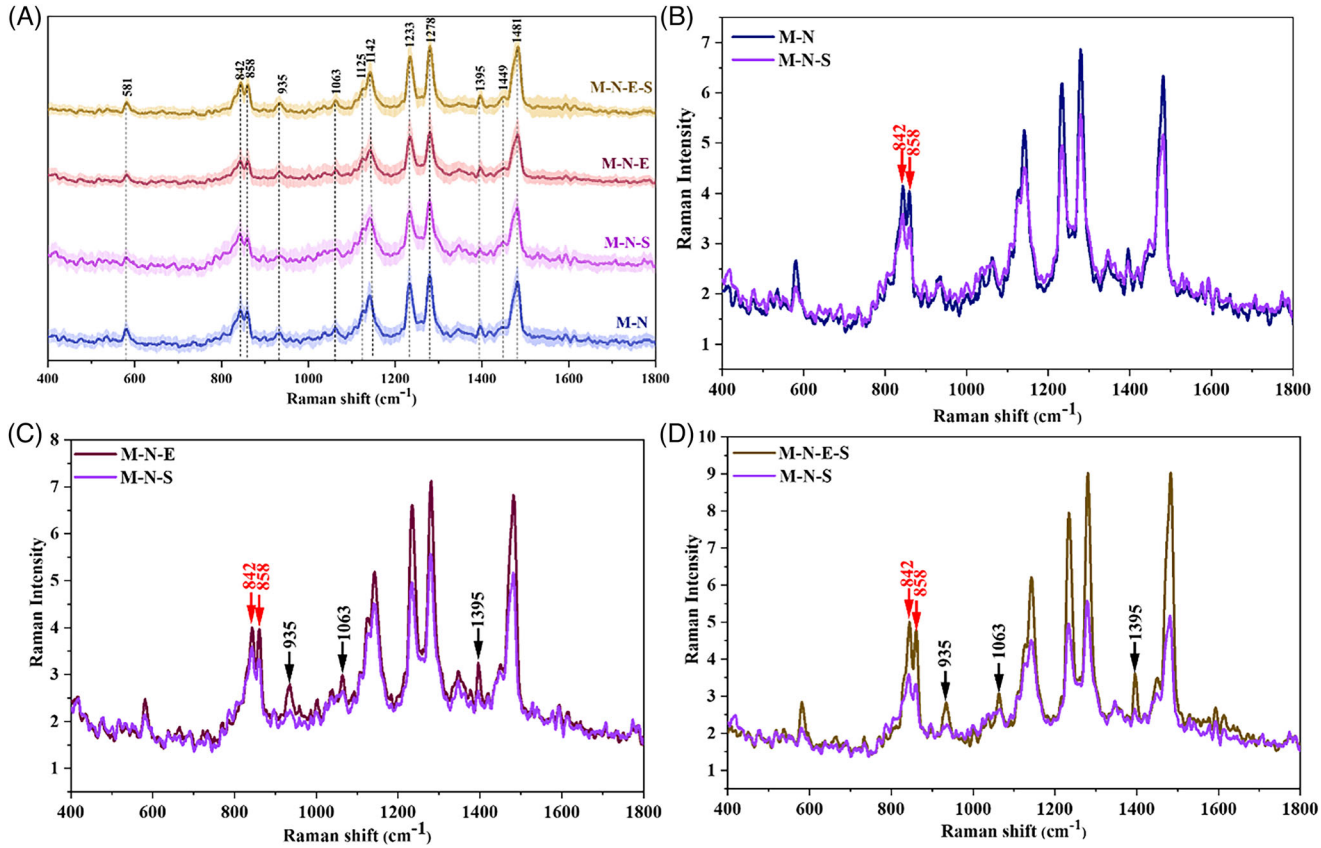


FIGURE 2 (A) Mean SERS spectra obtained SARS-CoV-2 VLP groups including M-N, M-N-S, M-N-E and M-N-E-S, (B–D) Comparison of M-N-S and other SARS-CoV-2 VLPs spectra. Red arrows indicate the presence of S protein and black arrows indicate peaks assigned to lipids.

N protein, which provides the basic capsid formation, and M protein are necessary structural components to formed VLP structures within the different groups of SARS-CoV-2 VLPs. The highest peak intensity in the spectra of M-N group was obtained at 858 cm^{-1} attributed to the NH_2, NH_2-terminal domain on the outside of the virus, which make the SARS-CoV-2 structure positive.[38]

The presence of S protein in the structure of SARS-CoV-2 VLPs is crucial during vaccine development since this protein plays an important role for entry of virus into the host cells. S protein is rich in aromatic groups where SARS-CoV-2 viruses interact with ACE2 receptor. Thus, a significant increase in the peak intensities were observed at 842 cm^{-1} (tyrosine), 1063, 1125 and 1142 cm^{-1} (phenylalanine, tryptophan) when the S protein was added to the M-N structure (M-N-S).[39,40] In addition, the peaks at 935, 1125,1395, and 1449 cm^{-1} on M-N-S spectra have the highest intensities among SARS-CoV-2 VLPs spectra. The strong affinity of disulfide bonds in the S protein to the AuNPs surface could be the reason of the increased the peak intensities. On the other hand, the addition of envelope protein to the M-N group decreased the intensity of the aforementioned peaks (842, 1063, 1125, and 1142 cm^{-1}) and significantly increased intensity of peaks associated with lipids (935 cm^{-1} (C–C stretch), 1063 cm^{-1} (C–C, C–O), 1395 cm^{-1} (C–H)). With the addition of the spike protein to the M-N-E group, the intensity of some peaks assigned for aromatic groups such as at 842, 858, 1233, 1278, and 1481 cm^{-1} increased. A peak at 1142 cm^{-1} was also raised, which could be related with the spike protein. Among the different SARS-CoV-2 VLPs groups, the ratio of peaks at 842 and 858 cm^{-1} attributed to the tyrosine and C-H or N-H deformation in the M-N structure, respectively, could be an indication of the presence of S protein in the SARS-CoV-2 VLP structure (Figure 2B). When we compared M-N and M-N-S groups, we observed that the 858/842 peak ratio became <1 when S protein was in the structure (M-N-S) or >1 in the absence of S protein in the M-N structure. When the spectra of M-N-S and M-N-E-S structures were compared, an increase in the peaks associated with lipids was observed, but the 858/842 ratio was <1 for the spectra of both groups (Figure 2D). The ratio of the 842 cm^{-1} peak to the 1233, 1278, and 1481 cm^{-1} peaks attributed to secondary structures of proteins, which are at >98%, >93%, >89.6% confidence levels, respectively, was similar results to those of 858/842 ratio, (Figure S9A–C).

The HIV-based VLP, which using HIV gag proteins produced by psPAX2 packaging plasmid, is different from the SARS-CoV-2 VLP structure. HIV-gag protein forms the capsid structure like N protein of SARS-CoV-2 and alone is sufficient to produce VLP. However, it should be noted that the SARS-CoV-2 N protein structure and HIV-gag protein structure are different.[41,42] This was reflected in the peak intensity rather than the peak position of the spectra. Especially, S protein in the HIV-based VLPs exhibited a different behavior compared to the SARS-CoV-2 VLPs. The comparison of the spectra

18607314, 2024, 1, Downloaded from https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/biot.202300191 by Koc University, Wiley Online Library on [15/02/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

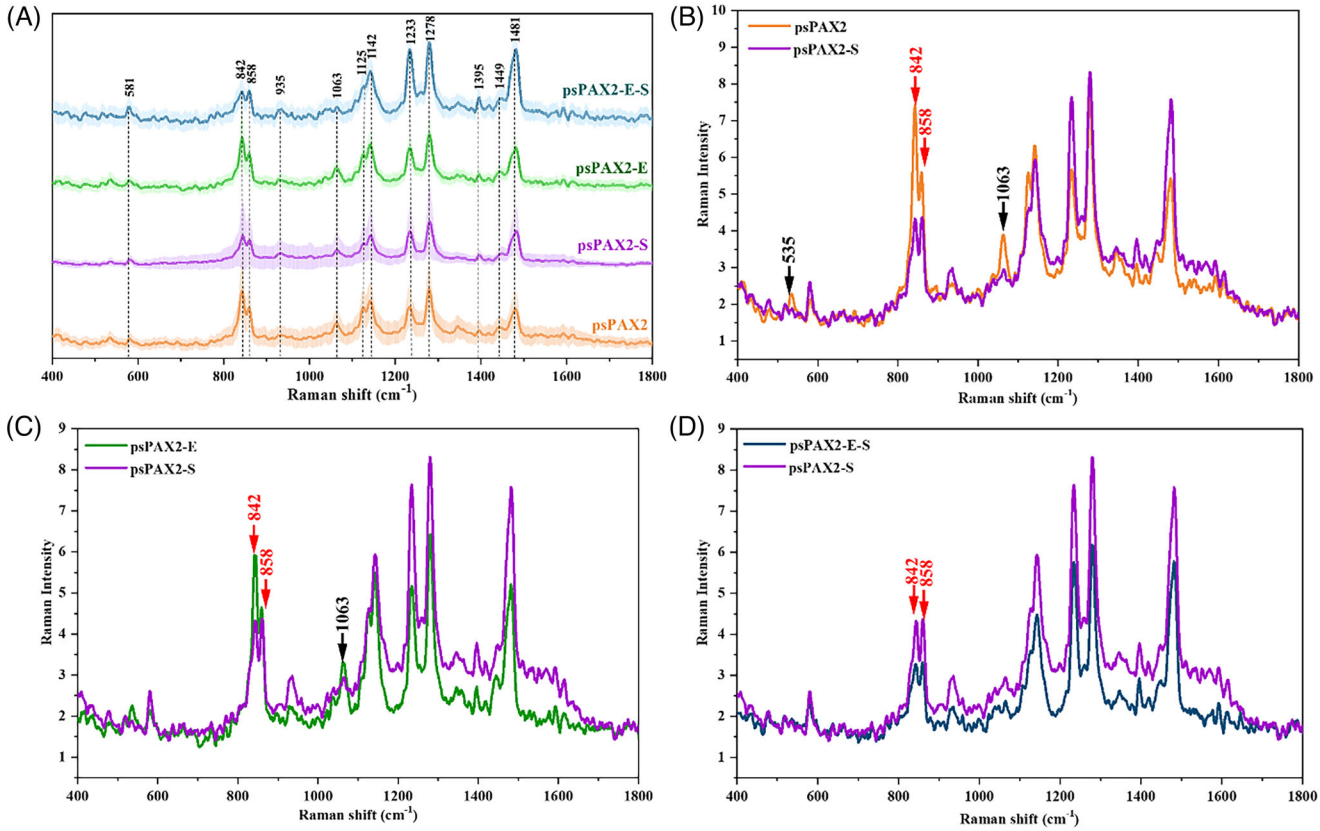


FIGURE 3 (A) SERS spectra obtained from HIV-based VLPs including psPAX2, psPAX2-S, psPAX2-E, and psPAX2-E-S. (B–D) Comparison of psPAX2-S and other HIV-based VLPs spectra. Red arrows indicate the presence of S protein and black arrows indicate peaks assigned to lipids and S–S bound.

of different groups of HIV-based VLPs is shown in Figure 3A. The psPAX2 structure generated clearly distinct peaks at 535, 842, 858, and 1063 cm^{-1} (disulfide (S-S), Tyrosine, C–H and C–C, C–O). With addition the envelope protein to psPAX2 structure (psPAX2-E) caused a significant increase of the peak intensity at 1063 cm^{-1} corresponding to the C–C and C–O stretching in lipid backbone.[43] When the spike protein is added to psPAX2 structure (psPAX2-S), an increase in the peak intensities at 535, 1233, 1272, and 1481 cm^{-1}, assigned to S-S, Amid III, and C-H, respectively was observed. Especially, the peak at 535 cm^{-1} could be due to disulfide bonds in S protein structure as it is mentioned in the literature.[44] Similar to SARS-CoV-2 VLPs, 858/842 peak ratio could be used to determine the presence of the S protein in the HIV-based VLP structure. However, we observed opposite results. As seen in Figure 3B–D, the peak intensity ratio became a value close to 1 when the structure included S protein. In contrast, the peak ratio became very much smaller in the psPAX2 and psPAX2-E structure. psPAX2 is an HIV-based lentiviral packaging system and has genetic information encoding tyrosine rich protein structures such as GaG, pol and gp41, revealing the tyrosine-enhanced peaks on the spectrum.[45] When spike protein added to psPAX2 (psPAX2-S) and psPAX2-E (psPAX2-E-S) structures, decreasing of these peak intensities may cause due to adsorption of the structures onto plasmonic surface through S-S an S protein structure.[46] Similar to SARS-CoV-2 VLPs ratio, the ratio of the 842/1233, 842/1278, and 842/1481, which are at >94.9%, >83.6%, >94.6% confidence levels, respectively, was similar results to the 858/842 ratio (Figure S9D-F-E).

Although the spectral differences revealed the presence of the S protein in the different SARS-CoV-2 VLP and HIV-based VLP structures, it is difficult to distinguish other significant compositional differences with the naked eye. Additionally, visual interpretation of the spectral differences may cause misinterpretation. Multivariate analysis techniques can discriminate between different SARS-CoV-2 VLP and HIV-based VLP groups containing different structural proteins of the virus by revealing small spectral differences.

In our study, we consistently incorporated M and N proteins in all VLP conditions, as they are crucial for SARS/SARS-2-based VLP production and release, as evidenced by prior research.[47–49] Although E and S proteins are less abundant, they integrate into the VLP structure when co-expressed.[48,49] One limitation of our characterization strategy is bulk analysis of samples. Techniques like NTA and TEM could give information on individual particles and elucidate size distribution and morphological variations of these particles. For instance, VLPs with the S protein should exhibit spike-like projections in EM, aligning with its distinct SERS peaks. Furthermore, while our emphasis was on collective SERS signals, single molecule SERS (sm-SERS) holds promise for in-depth analysis of individual VLP compositions, offering insights into sample homogeneity. Optimizing our SERS substrate could make

18607314, 2024, 1, Downloaded from https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/biot.202300191 by Koc University, Wiley Online Library on [15/02/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 1 Accuracy values of classification techniques for SARS-CoV-2 VLPs and HIV-based VLPs.

Classifier	Samples	Accuracy [%]
kNN	HIV-based VLPs	55
	SARS-CoV-2 VLPs	62.5
RF	HIV-based VLPs	85
	SARS-CoV-2 VLPs	72.5
SVM (poly)	HIV-based VLPs	90
	SARS-CoV-2 VLPs	87.5
SVM (Gaussian)	HIV-based VLPs	85
	SARS-CoV-2 VLPs	77.5
SVM (Linear)	HIV-based VLPs	92.5
	SARS-CoV-2 VLPs	85

sm-SERS an instrumental method for evaluating VLP quality on a single particle level.

3.4 | Multivariate analysis

For the classification of SARS-CoV-2 VLPs and HIV-based VLPs, SVM, RF, and kNN were used to discriminate and define unique spectral profiles of both SARS-CoV-2 VLPs and HIV-based VLPs. These classifiers have distinct benefits and drawbacks, these classifiers can be used for both regression and classification.[50] The performance of these classifiers has not been tested for COVID-19 detection studies with Raman spectroscopy or SERS.[37,51,52]

SVM is a model that generates optimal hyperplanes or decision boundaries and separates different classes in a high-dimensional space. SVM can also deal with nonlinear situations where boundaries are very complex due to the complexity of biological spectra. There is a kernel trick that allows SVMs to define complex boundaries.[53] The kernels can be linear, polynomial, Gaussian, or sigmoid and vary according to dataset. For example, the polynomial kernel is a non-stationary kernel and is suitable for problems where all training data is normalized.[54] The kNN classifies according to k value of the class (learning set) and the nearest neighbor where the sample data point is located.[55] While there is no training for kNN, it has disadvantages such as requiring memory and being costly with increasing dataset and feature size. Moreover, RF is a model which classifies by generating multiple independent decision trees.[56]

In this study, the dataset for SARS-CoV-2 VLPs consists of 280 spectra (70 spectra for each group) and the dataset for HIV-based VLPs includes 280 spectra (70 spectra for each group). Two hundred and forty spectra were used for training, and 40 spectra were used to test the spectra of two viral structures. It is worth nothing that data was split randomly. The accuracy of each classifier was calculated and demonstrated in Table 1.

According to these analyses, the poly SVM classifier showed the highest classification accuracy with 87.5% and the lowest accuracy value was 62.5% in the kNN classifier for SARS-CoV-2 VLPs. For HIV-based VLPs, the linear SVM classifier showed the highest level of accuracy with 92.5%, while the kNN classifier had the lowest classification accuracy with 55%. RF, Gaussian SVM, and poly SVM demonstrated lower results than linear SVM classifiers with orderly accuracy values of 85%, 85%, and 90%, respectively. The accuracy value was obtained with a confusion matrix. Figure S10A demonstrates the confusion matrix of poly SVM for SARS-CoV-2 VLPs which provides the best accuracy value among others. All M-N and M-N-E samples were correctly classified except four spectra from the M-N-S sample misclassified as M-N and 1 spectrum from the M-N-E-S sample identified as M-N-E. Figure S10B shows the confusion matrix of linear SVM for HIV-based VLPs which has the best accuracy among others. Here, all psPAX2-S, psPAX2-E, and psPAX2-E-S samples were correctly identified, while three spectra from the psPAX2 sample were misclassified as psPAX2-S.

Indeed, there some studies proved the superiority of SVM over the other algorithms in viral and Raman spectra classification.[57–60] In our study, we used the "kernel trick" for SVM, which is a clever mathematical technique that allows SVM to handle nonlinear situations. The kernel functions, such as linear, polynomial, Gaussian, and sigmoid, transform the original data into a higher-dimensional space. This transformation enables SVM to find complex decision boundaries in the feature space, even when the relationships between data points are not linear.[53] In the case of Raman spectroscopy data, the molecular spectra of SARS-CoV-2 VLPs and HIV-based VLPs can be intricate and non-linear. By applying the appropriate kernel function, SVM can effectively map the Raman spectra into a higher-dimensional space, where it can identify more intricate patterns and relationships between different viral structures. This ability to define complex boundaries in feature space is crucial in accurately classifying the different viral structures and achieving superior accuracy compared to other classifiers.

In addition, the confusion matrices of the other classifiers (kNN and RF) showed some misclassifications in our study. SVM outperformed these classifiers, especially in cases of higher accuracy. The misclassifications observed in other classifiers could be attributed to their limitations in capturing specific characteristics of the Raman spectra. For example, kNN classifies data points based on the majority class of its k-nearest neighbors. In Raman spectroscopy, the kNN algorithm might face challenges when nearest neighbors are not representative enough to identify subtle differences between the viral structures. As a result, some spectra may be misclassified. Similarly, RF generates multiple independent decision trees, and the final classification is based on a majority vote from the trees. The misclassifications observed with RF might arise due to different decisions made by individual trees, which can lead to inconsistencies in the final classification.

4 | CONCLUSION

SERS combined with machine learning presents a rapid, cost effective and reliable detection and discrimination for viral structures. This method can offer a powerful control over pandemics for prevention

18607314, 2024, 1, Downloaded from https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/biot.202300191 by Koc University, Wiley Online Library on [15/02/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

spread of pandemic and characterize viral structures for development of novel therapeutics. In this study, we demonstrated a SERS-based approach for detection of COVID-19 and characterization of virus-like particles by revealing precious molecular level information of different structures of VLP. Combined with machine learning techniques, this tool showed promising results for the classifying SERS spectra of different VLP structures. This approach can be helpful for further detection of the rapid and emerging mutations of the pandemics, which will be helpful not only for detection but also for the finding and characterization of novel vaccine and other VLP-based therapeutics.

AUTHOR CONTRIBUTIONS

Munevver Akdeniz: Conceptualization; Formal analysis; Investigation; Methodology; Software; Validation; Visualization; Writing – original draft. **Zakarya Al-shaebi**: Formal analysis; Methodology; Software; Writing – original draft. **Mine Altunbek**: Methodology; Validation; Writing – review & editing. **Canan Bayraktar**: Methodology; Resources; Writing – original draft. **Alisan Kayabolen**: Investigation; Methodology; Resources; Writing – review & editing. **Tugba Bagci-Onder**: Funding acquisition; Supervision; Writing – review & editing. **Omer Aydin**: Conceptualization; Funding acquisition; Methodology; Project administration; Resources; Supervision; Writing – review & editing

ACKNOWLEDGMENTS

This work was supported by the Erciyes University Scientific Research Projects Coordination Unit under grant numbers, # FYL-2021-10780.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Omer Aydin https://orcid.org/0000-0002-9028-8786

REFERENCES

1. Hu, B., Guo, H., Zhou, P., & Shi, Z.-L. (2021). Characteristics of SARS-CoV-2 and COVID-19. *Nature Reviews Microbiology*, *19*, 141–154.
2. Mariano, G., Farthing, R. J., Lale-Farjat, S. L., & Bergeron, J. R. (2020). Structural characterization of SARS-CoV-2: Where we are, and where we need to be. *Frontiers in Molecular Biosciences*, *7*, 605236.
3. Satarker, S., & Nampoothiri, M. (2020). Structural proteins in severe acute respiratory syndrome coronavirus-2. *Archives of Medical Research*, *51*, 482–491.
4. Grgacic, E. V. L., & Anderson, D. A. (2006). Virus-like particles: Passport to immune recognition. *Methods (San Diego, Calif.)*, *40*, 60–65.
5. Shirbaghaee, Z., & Bolhassani, A. (2016). Different applications of virus-like particles in biology and medicine: Vaccination and delivery systems. *Biopolymers*, *105*, 113–132.
6. Nooraei, S., Bahrulolum, H., Hoseini, Z. S., Katalani, C., Hajizade, A., Easton, A. J., & Ahmadian, G. (2021). Virus-like particles: Preparation, immunogenicity and their roles as nanovaccines and drug nanocarriers. *Journal of Nanbiotechnology*, *19*, 59.
7. Yilmaz, I. C., Ipekoglu, E. M., Bulbul, A., Turay, N., Yildirim, M., Evcili, I., Yilmaz, N. S., Guvencli, N., Aydin, Y., Gungor, B., Saraydar, B., Bartan, A. G., Ibibik, B., Bildik, T., Baydemir, İ., Sanli, H. A., Kayaoglu, B., Ceylan, Y., Yildirim, T., … Gursel, M. (2021). Development and preclinical evaluation of virus-like particle vaccine against COVID-19 infection. *Allergy*, *77*(1), 258–270. *n/a*.
8. Lua, L. H. L., Connors, N. K., Sainsbury, F., Chuan, Y. P., Wibowo, N., & Middelberg, A. P. J. (2014). Bioengineering virus-like particles as vaccines. *Biotechnology and Bioengineering*, *111*, 425–440.
9. Xu, R., Shi, M., Li, J., Song, P., & Li, N. (2020). Construction of SARS-CoV-2 Virus-Like Particles by Mammalian Expression System. Frontiers in Bioengineering and Biotechnology, *8*, 862.
10. Carvalho, S. B., Silva, R. J. S., Sousa, M. F. Q., Peixoto, C., Roldão, A., Carrondo, M. J. T., & Alves, P. M. (2022). Bioanalytics for influenza virus-like particle characterization and process monitoring. Frontiers in Bioengineering and Biotechnology, *10*, 805176.
11. Jalandra, R., Yadav, A. K., Verma, D., Dalal, N., Sharma, M., Singh, R., Kumar, A., & Solanki, P. R. (2020). Strategies and perspectives to develop SARS-CoV-2 detection methods and diagnostics. *Biomedicine & Pharmacotherapy*, *129*, 110446.
12. Eftekhari, A., Alipour, M., Chodari, L., Maleki Dizaj, S., Ardalan, M., Samiei, M., Sharifi, S., Zununi Vahed, S., Huseynova, I., Khalilov, R., Ahmadian, E., & Cucchiarini, M. (2021). A Comprehensive Review of Detection Methods for SARS-CoV-2. *Microorganisms*, *9*(2), 232.
13. Stöckel, S., Kirchhoff, J., Neugebauer, U., Rösch, P., & Popp, J. (2016). The application of Raman spectroscopy for the detection and identification of microorganisms. *Journal of Raman Specroscopy*, *47*, 89–109.
14. Pezzotti, G. (2021). Raman spectroscopy in cell biology and microbiology. *Journal of Raman Specroscopy*, *52*, 2348–2443.
15. Langer, J., Jimenez De Aberasturi, D., Aizpurua, J., Alvarez-Puebla, R. A., Auguié, B., Baumberg, J. J., Bazan, G. C., Bell, S. E. J., Boisen, A., Brolo, A. G., Choo, J., Cialla-May, D., Deckert, V., Fabris, L., Faulds, K., García De Abajo, F. J., Goodacre, R., Graham, D., Haes, A. J., … Liz-Marzán, L. M. (2020). Present and Future of Surface-Enhanced Raman Scattering. *ACS Nano*, *14*, 28–117.
16. Panikar, S. S., Cialla-May, D., De La Rosa, E., Salas, P., & Popp, J. (2021). Towards translation of surface-enhanced Raman spectroscopy (SERS) to clinical practice: Progress and trends. *TrAC Trends in Analytical Chemistry*, *134*, 116122.
17. Sharma, B., Fernanda Cardinal, M., Kleinman, S. L., Greeneltch, N. G., Frontiera, R. R., Blaber, M. G., Schatz, G. C., & Van Duyne, R. P. (2013). High-performance SERS substrates: Advances and challenges. *MRS Bulletin*, *38*, 615–624.
18. Kahraman, M., Aydin, Ö., & Çulha, M. (2009). Size effect of 3D aggregates assembled from silver nanoparticles on surface-enhanced Raman scattering. *Chemphyschem*, *10*, 537–542.
19. Zhu, M., Li, M., Su, M., Liu, J., Liu, B., Ge, Y., Liu, H., & Hu, J. (2021). Can "Hot Spots" be stable enough for surface-enhanced Raman scattering? *Journal of Physical Chemistry C*, *125*, 13443–13448.
20. Mehigan, S., Smyth, C. A., & Mccabe, E. M. (2015). Bridging the gap between SERS enhancement and reproducibility by salt aggregated silver nanoparticles. *Nanomaterials and Nanotechnology*, *5*, 5.
21. Chan, M. Y., Leng, W., & Vikesland, P. J. (2018). Surface-enhanced raman spectroscopy characterization of salt-induced aggregation of gold nanoparticles. *Chemphyschem*, *19*, 24–28.
22. Lim, J.-Y., Nam, J.-S., Yang, S.-E., Shin, H., Jang, Y.-H., Bae, G.-U., Kang, T., Lim, K.-I., & Choi, Y. (2015). Identification of newly emerging influenza viruses by surface-enhanced Raman spectroscopy. *Analytical Chemistry*, *87*, 11652–11659.
23. Akdeniz, M., Uysal Ciloglu, F., Tunc, C. U., Yilmaz, U., Kanarya, D., Atalay, P., & Aydin, O. (2022). Investigation of mammalian cells expressing SARS-CoV-2 proteins by surface-enhanced Raman scattering and multivariate analysis. *Analyst*, 1213–1221.
24. Bell, S. E. J., & Sirimuthu, N. M. S. (2006). Surface-Enhanced Raman Spectroscopy (SERS) for sub-micromolar detection of DNA/RNA

18607314, 2024, 1, Downloaded from https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/biot.202300191 by Koc University, Wiley Online Library on [15/02/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

mononucleotides. *Journal of the American Chemical Society*, *128*, 15580–15581.

25. Sitjar, J., Xu, H.-Z., Liu, C.-Y., Wang, J.-R., Liao, J.-D., Tsai, H.-P., Lee, H., Liu, B. H., & Chang, C.-W. (2022). Synergistic surface-enhanced Raman scattering effect to distinguish live SARS-CoV-2 S pseudovirus. *Analytica Chimica Acta*, *1193*, 339406–339406.
26. Lebedev, N., Griva, I., Dressick, W. J., Phelps, J., Johnson, J. E., Meshcheriakova, Y., Lomonossoff, G. P., & Soto, C. M. (2016). A virus-based nanoplasmonic structure as a surface-enhanced Raman biosensor. *Biosensors & Bioelectronics*, *77*, 306–314.
27. Sitjar, J., Liao, J.-D., Lee, H., Tsai, H.-P., Wang, J.-R., & Liu, P.-Y. (2021). Challenges of SERS technology as a non-nucleic acid or -antigen detection method for SARS-CoV-2 virus and its variants. *Biosensors & Bioelectronics*, *181*, 113153.
28. Ralbovsky, N. M., & Lednev, I. K. (2020). Towards development of a novel universal medical diagnostic method: Raman spectroscopy and machine learning. *Chemical Society Reviews*, *49*, 7428–7453.
29. Uysal Ciloglu, F., Saridag, A. M., Kilic, I. H., Tokmakci, M., Kahraman, M., & Aydin, O. (2020). Identification of methicillin-resistant *Staphylococcus aureus* bacteria using surface-enhanced Raman spectroscopy and machine learning techniques. *Analyst*, *145*, 7559–7570.
30. Ciloglu, F. U., Hora, M., Gundogdu, A., Kahraman, M., Tokmakci, M., & Aydin, O. (2022). SERS-based sensor with a machine learning based effective feature extraction technique for fast detection of colistin-resistant *Klebsiella pneumoniae*. *Analytica Chimica Acta*, *1221*, 340094.
31. Bai, X., Lin, J., Wu, X., Lin, Y., Zhao, X., Du, W., Gao, J., Hu, Z., Xu, Q., Li, T., & Yu, Y. (2022). Label-free detection of bladder cancer and kidney cancer plasma based on SERS and multivariate statistical algorithm. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, *279*, 121336.
32. Turkevich, J., Stevenson, P. C., & Hillier, J. (1951). A study of the nucleation and growth processes in the synthesis of colloidal gold. *Discussions of the Faraday Society*, *11*, 55–75.
33. Lim, J.-Y., Nam, J.-S., Shin, H., Park, J., Song, H.-I., Kang, M., Lim, K.-I., & Choi, Y. (2019). Identification of newly emerging influenza viruses by detecting the virally infected cells based on surface enhanced Raman spectroscopy and principal component analysis. *Analytical Chemistry*, *91*, 5677–5684.
34. Hong, S., & Li, X. (2013). Optimal size of gold nanoparticles for surface-enhanced Raman spectroscopy under different conditions. *Journal of Nanomaterials*, 790323.
35. Altunbek, M., Kuku, G., & Culha, M. (2016). Gold nanoparticles in single-cell analysis for surface enhanced Raman scattering. *Molecules (Basel, Switzerland)*, *21*, 1617.
36. Karimi, S., Moshaii, A., & Nikkhah, M. (2019). Controlled synthesis of colloidal monodisperse gold nanoparticles in a wide range of sizes; investigating the effect of reducing agent. *Materials Research Express*, *6*, 1150f2.
37. Sanchez, J. E., Jaramillo, S. A., Settles, E., Velazquez Salazar, J. J., Lehr, A., Gonzalez, J., Rodríguez Aranda, C., Navarro-Contreras, H. R., Raniere, M. O., Harvey, M., Wagner, D. M., Koppisch, A., Kellar, R., Keim, P., & Jose Yacaman, M. (2021). Detection of SARS-CoV-2 and its S and N proteins using surface enhanced Raman spectroscopy. *RSC Advances*, *11*, 25788–25794.
38. Thomas, S. (2020). The structure of the membrane protein of SARS-CoV-2 resembles the sugar transporter SemiSWEET. *Pathogens and Immunity*, *5*, 342.
39. Du, L., He, Y., Zhou, Y., Liu, S., Zheng, B. O.-J., & Jiang, S. (2009). The spike protein of SARS-CoV–a target for vaccine and therapeutic development. *Nature Reviews Microbiology*, *7*, 226–236.
40. Paria, D., Kwok, K. S., Raj, P., Zheng, P., Gracias, D. H., & Barman, I. (2022). Label-free spectroscopic SARS-CoV-2 detection on versatile nanoimprinted substrates. *Nano Letters*, *22*, 3620–3627.
41. Ye, Q., West, A. M. V., Silletti, S., & Corbett, K. D. (2020). Architecture and self-assembly of the SARS-CoV-2 nucleocapsid protein. *Protein Science*, *29*, 1890–1901.
42. Bell, N. M., & Lever, A. M. L. (2013). HIV Gag polyprotein: Processing and early viral particle assembly. *Trends in Microbiology*, *21*, 136–144.
43. Schoeman, D., & Fielding, B. C. (2019). Coronavirus envelope protein: Current knowledge. *Virology Journal*, *16*, 69.
44. Wu, Q., Zhang, Y., Lu, H., Wang, J., He, X., Liu, Y., Ye, C., Lin, W., Hu, J., Ji, J., Xu, J., Ye, J., Hu, Y., Chen, W., Li, S., Wang, J., Wang, J., Bi, S., & Yang, H. (2003). The E protein is a multifunctional membrane protein of SARS-CoV. *Genomics, Proteomics & Bioinformatics*, *1*, 131–144.
45. Chen, M., & Zhang, X.-E. (2021). Construction and applications of SARS-CoV-2 pseudoviruses: A mini review. *International Journal of Biological Sciences*, *17*, 1574–1580.
46. Gupta, D., Sharma, P., Singh, M., Kumar, M., Ethayathulla, A. S., & Kaur, P. (2021). Structural and functional insights into the spike protein mutations of emerging SARS-CoV-2 variants. *Cellular and Molecular Life Sciences*, *78*, 7967–7989.
47. Huang, Y., Yang, Z.-Y., Kong, W.-P., & Nabel, G. J. (2004). Generation of synthetic severe acute respiratory syndrome coronavirus pseudoparticles: Implications for assembly and vaccine production. *Journal of Virology*, *78*, 12557–12565.
48. Plescia, C. B., David, E. A., Patra, D., Sengupta, R., Amiar, S., Su, Y., & Stahelin, R. V. (2021). SARS-CoV-2 viral budding and entry can be modeled using BSL-2 level virus-like particles. *The Journal of Biological Chemistry*, *296*, 100103.
49. Gourdelier, M., Swain, J., Arone, C., Mouttou, A., Bracquemond, D., Merida, P., Saffarian, S., Lyonnais, S., Favard, C., & Muriaux, D. (2022). Optimized production and fluorescent labeling of SARS-CoV-2 virus-like particles. *Scientific Reports*, *12*, 14651.
50. Lussier, F., Thibault, V., Charron, B., Wallace, G. Q., & Masson, J.-F. (2020). Deep learning and artificial intelligence methods for Raman and surface-enhanced Raman scattering. *TrAC Trends in Analytical Chemistry*, *124*, 115796.
51. Carlomagno, C., Bertazioli, D., Gualerzi, A., Picciolini, S., Banfi, P. I., Lax, A., Messina, E., Navarro, J., Bianchi, L., Caronni, A., Marenco, F., Monteleone, S., Arienti, C., & Bedoni, M. (2021). COVID-19 salivary Raman fingerprint: Innovative approach for the detection of current and past SARS-CoV-2 infections. *Scientific Reports*, *11*, 4943.
52. Huang, J., Wen, J., Zhou, M., Ni, S., Le, W., Chen, G., Wei, L., Zeng, Y., Qi, D., Pan, M., Xu, J., Wu, Y., Li, Z., Feng, Y., Zhao, Z., He, Z., Li, B., Zhao, S., Zhang, B., … Du, K. (2021). On-site detection of SARS-CoV-2 antigen by deep learning-based surface-enhanced raman spectroscopy and its biochemical foundations. *Analytical Chemistry*, *93*, 9174–9182.
53. Xu, Y., Zomer, S., & Brereton, R. G. (2006). Support vector machines: A recent method for classification in chemometrics. *Critical Reviews in Analytical Chemistry*, *36*, 177–188.
54. Patle, A., & Chouhan, D. S. (2013). 2013 International Conference on Advances in Technology and Engineering (ICATE), 1–9.
55. Zhang, S. (2020). Cost-sensitive KNN classification. *Neurocomputing*, *391*, 234–242.
56. Pan, L., Zhang, P., Daengngam, C., Peng, S., & Chongcheawchamnan, M. (2022). A review of artificial intelligence methods combined with Raman spectroscopy to identify the composition of substances. *Journal of Raman Specroscopy*, *53*, 6–19.
57. Ciloglu, F. U., Caliskan, A., Saridag, A. M., Kilic, I. H., Tokmakci, M., Kahraman, M., & Aydin, O. (2021). Drug-resistant *Staphylococcus aureus* bacteria detection by combining surface-enhanced Raman spectroscopy (SERS) and deep learning techniques. *Scientific Reports*, *11*, 18444.
58. Zheng, Q., Li, J., Yang, L., Zheng, B., Wang, J., Lv, N., Luo, J., Martin, F. L., Liu, D., & He, J. (2020). Raman spectroscopy as a potential diagnostic tool to analyse biochemical alterations in lung cancer. *Analyst*, *145*, 385–392.

18607314, 2024, 1, Downloaded from https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/bit.28230019] by Koc University, Wiley Online Library on [15/02/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

59. Radzol, A. R. M., Lee, K. Y., Mansor, W., & Omar, I. S. (2016). 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). 6206–6209.
60. Khan, S., Ullah, R., Khan, A., Ashraf, R., Ali, H., Bilal, M., & Saleem, M. (2018). Analysis of hepatitis B virus infection in blood sera using Raman spectroscopy and machine learning. *Photodiagnosis and Photodynamic Therapy*, *23*, 89–93.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Akdeniz, M., Al-Shaebi, Z., Altunbek, M., Bayraktar, C., Kayabolen, A., Bagci-Onder, T., & Aydin, O. (2024). Characterization and discrimination of spike protein in SARS-CoV-2 virus-like particles via surface-enhanced Raman spectroscopy. *Biotechnology Journal*, *19*, e2300191. https://doi.org/10.1002/biot.202300191

18607314, 2024, 1, Downloaded from https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/biot.202300191 by Koc University, Wiley Online Library on [15/02/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License