ACS Photonics

pubs.acs.org/journal/apchd5

Article

Understanding the Excitation Wavelength Dependence and Thermal Stability of the SARS-CoV-2 Receptor-Binding Domain Using Surface-Enhanced Raman Scattering and Machine Learning

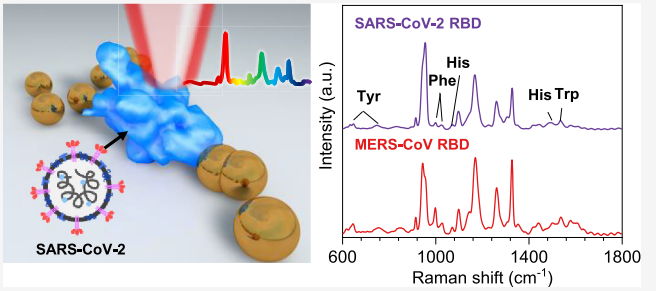
Kunyan Zhang, Ziyang Wang, He Liu, Néstor Perea-López, Jeewan C. Ranasinghe, George Bepete, Allen M. Minns, Randall M. Rossi, Scott E. Lindner, Sharon X. Huang, Mauricio Terrones,* and Shengxi Huang*

Cite This: https://doi.org/10.1021/acsphotonics.2c00456

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: COVID-19 has cost millions of lives worldwide. The constant mutation of SARS-CoV-2 calls for thorough research to facilitate the development of variant surveillance. In this work, we studied the fundamental properties related to the optical identification of the receptor-binding domain (RBD) of SARS-CoV-2 spike protein, a key component of viral infection. The Raman modes of the SARS-CoV-2 RBD were captured by surface-enhanced Raman spectroscopy (SERS) using gold nanoparticles (AuNPs). The observed Raman enhancement strongly depends on the excitation wavelength as a result of the aggregation of AuNPs. The characteristic Raman spectra of RBDs from SARS-CoV-2 and MERS-CoV were analyzed by principal component analysis that reveals the role of secondary structures in the SERS process, which is corroborated with the thermal stability under laser heating. We can easily distinguish the Raman spectra of two RBDs using machine learning algorithms with accuracy, precision, recall, and F_1 scores all over 95%. Our work provides an in-depth understanding of the SARS-CoV-2 RBD and paves the way toward rapid analysis and discrimination of complex proteins of infectious viruses and other biomolecules.



KEYWORDS: *COVID-19, amino acid, secondary structure, nanoparticle, Raman spectroscopy, machine learning*

1. INTRODUCTION

In early 2020, COVID-19 rapidly became the leading cause of death and the exponential increase in positive cases overwhelmed substantial portions of the healthcare system.[1] Although highly protective vaccines have been developed,[2−4] mutations of the genome of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) continue to generate new variants. Particularly, amino acid changes in the receptor-binding domain (RBD) of SARS-CoV-2 spike protein, an essential component for binding to human angiotensin-converting enzyme-2 (ACE2) receptors,[5] can partially escape vaccine protection and hinder the effort of rapid detection and contact tracing.[6−10] For example, the Delta variant that emerged in late 2020 has amino acid changes in the RBD of spike protein that leads to reduced vaccine efficacy.[11] The more recent variant of concern designated in November 2021, Omicron, harbors amino acid substitutions at around 30 locations in the spike protein, around half of which are in the RBD.[12] Therefore, structural identification of the RBD is a critical factor to enable variant monitoring based on the amino acid substitutions and large-scale testing targeted at the SARS-CoV-2 virus that is constantly mutating and evolving.

The current COVID-19 testing methods, including nucleotide detection and antigen detection-based rapid tests and immunoassays, are relatively time-consuming. Most testing methods demand analyses at specialized centers by skilled staff and require prior knowledge of the genomic sequence of either the strain or the corresponding antibodies. In contrast, Raman spectroscopy, a rapid and label-free optical characterization technique, is an alternative to effectively identify different strains and subtypes of infectious viruses.[13] This technique also provides fundamental information about the optical properties of analytes by measuring the chemical bond vibrations caused by light-matter interactions.[13] However, Raman scattering of biological samples suffers from a low scattering efficiency of 10^{-6} to 10^{-10}, requiring enhancement methods to amplify the signal,[14,15] one of which being surface-enhanced Raman

Received: March 23, 2022

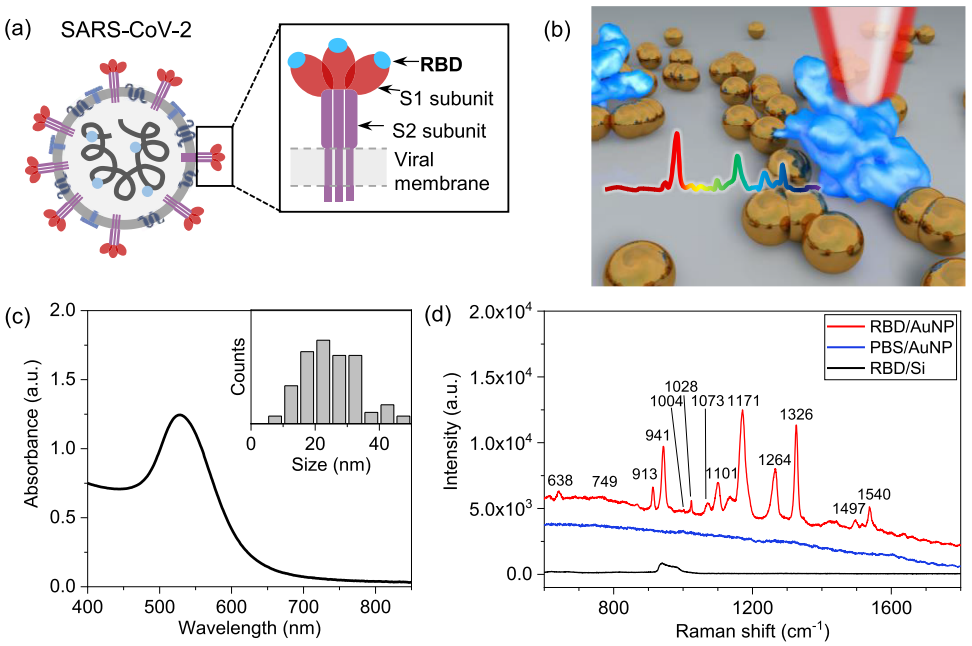


Figure 1. SERS of SARS-CoV-2 RBD. (a) Illustration of SARS-CoV-2. The inset shows the spike protein consisting of S1 and S2 subunits. SARS-CoV-2 RBD is in the S1 subunit. (b) Schematics of SERS of SARS-CoV-2 RBD on AuNPs. (c) UV−vis spectroscopy of AuNPs in water. The absorption peak is at 528 nm. The inset shows the size distribution of AuNPs. (d) Raman spectra of SARS-CoV-2 RBD on AuNPs, PBD solution on AuNPs, and SARS-CoV-2 RBD on silicon measured using a 785 nm laser.

spectroscopy (SERS). SERS utilizes substrates ranging from metal nanoparticles,[16] metamaterials,[17] to two-dimensional materials[18−21] to increase the scattering efficiencies of different analytes. With the aid of machine learning, SERS is capable of identifying the important virus component that contributes to the Raman signals and distinguishing between different types of influenza viruses.[13,22] Although optical techniques, such as Raman spectroscopy, have been applied to identify SARS-CoV-2 and its protein components,[23−25] there have been scarce investigations of their fundamental properties including excitation wavelength dependence and thermal stability, which determine the SERS detection efficacy.

In this work, we studied the RBD of SARS-CoV-2 spike protein, a key component for viral infection, using SERS and machine learning. Several Raman vibrational modes of the SARS-CoV-2 RBD were observed from 600 to 1800 cm^{-1} using colloidal gold nanoparticles as the SERS substrate. The Raman signals are dependent on the laser wavelength, in which the 785 nm laser yields a stronger Raman response than 532 and 633 nm lasers for the SARS-CoV-2 RBD, as well as for the RBD of another infectious coronavirus, Middle East respiratory syndrome coronavirus (MERS-CoV). The strong dependence of SERS performance on the excitation wavelength is explained by the red-shift of plasmon resonance as a consequence of nanoparticle aggregation. We also examined the thermal stability of two types of the RBD to improve the practicality of SERS. Moreover, our principal component analysis of the Raman spectra reveals the relationship between the nanostructure of the RBD and SERS responses. Since the Raman fingerprints of RBDs reflect the structural and compositional characteristics, the two types of RBDs can be distinguished using machine learning classifiers with accuracy, precision, recall, and F_1 scores of over 95%. The in-depth analysis of the feature weights based on linear support vector machine also indicates the contribution from secondary structures and Raman-active amino acids including tyrosine, phenylalanine, and histidine to the SERS spectra, consistent with the sequencing results. Our investigation of the fundamental optical properties of the SARS-CoV-2 RBD and the underlying SERS mechanism could facilitate real-time monitoring of the constantly changing mutants of the circulating SARS-CoV-2.

2. RESULTS AND DISCUSSION

2.1. SERS of SARS-CoV-2 RBD. The RBD of SARS-CoV-2 is a key component of the S1 subunit of spike glycoprotein (Figure 1a). It enables the binding to the host ACE2 receptor and allows membrane fusion with host cells to initiate viral infection.[5] Moreover, changes in the RBD enable the transmission of SARS-CoV-2 between different hosts.[26,27] Thus, the RBD is a key target in developing vaccines and neutralizing antibodies. Detailed expression of the purified SARS-CoV-2 RBD protein used in this work is described in the Methods section. For the SERS measurement, colloidal gold nanoparticles (AuNPs) with an average diameter of 20 nm were deposited on silicon and used as the SERS substrates (Figure 1b). The 20 nm AuNPs display an optical absorption at 528 nm as shown by ultraviolet−visible (UV−vis) spectroscopy in Figure 1c. The highly purified RBD in phosphate-buffered saline (PBS) was then deposited on the SERS substrate and dried before SERS measurements. As shown in Figure 1d, SARS-CoV-2 RBD on 20 nm AuNPs (RBD/AuNP) shows distinct vibrational modes at the Raman frequencies of 913, 941, 1101, 1133, 1171, 1264, and 1326 cm^{-1}. These Raman modes are otherwise absent for PBS solution deposited on AuNPs (PBS/AuNP) and SARS-CoV-2 RBD deposited on bare silicon (RBD/Si) under a 785 nm laser excitation.

The distinct Raman modes observed in the Raman spectra result from secondary structures and specific amino acids that compose the SARS-CoV-2 RBD as shown in Table 1. The strong Raman modes at 1264 and 1326 cm^{-1} are assigned to

https://doi.org/10.1021/acsphotonics.2c00456
ACS Photonics XXXX, XXX, XXX−XXX

Table 1. Assignment of the Raman Peaks of SARS-CoV-2 RBD

Raman modes (cm^{-1})	assignment
638	tyrosine (α-helix)
749	tyrosine (α-helix)
913	deformation vibrations of CH_2
941	stretching vibrations of $C^{\alpha}-C^{\beta}$ (α-helix)
1004	phenylalanine
1028	phenylalanine
1073	histidine
1101	stretching vibrations of C−C and $C^{\alpha}-N$
1133	stretching vibrations of C−C and C−N
1171	deformation vibrations of CH_3
1264	amide III (α-helix)
1326	amide III
1497	histidine
1540	tryptophan

the polypeptide backbone (amide III), and the peaks located at 913 and 1171 cm^{-1} are attributed to the deformation vibrations of CH_2 and CH_3, while the modes at 941, 1101, and 1133 cm^{-1} originate from the stretching vibrations of C−C and C−N.[28] The amide III bands of the RBD of SARS-CoV-2 are in agreement with frequencies of the spike protein as well as the deactivated SARS-CoV-2 virion particle.[29] The amide I band is not consistently observed in the Raman spectrum of SARS-CoV-2 RBD, which is explained by the size of the amino acid side chain that prevents immediate contact between the peptide and the AuNP.[30] The absence of amide I has also been reported for SERS of SARS-CoV-2 virion particles using Au−Cu nanostars and Raman measurement of SARS-CoV-2 in saliva.[24,29] In contrast, specific amino acids lead to less-prominent peaks in the Raman spectra. The Raman modes positioned at 638 and 749 cm^{-1}, 1004 and 1028 cm^{-1}, 1073 and 1497 cm^{-1}, and 1540 cm^{-1} have contributions from tyrosine, phenylalanine, histidine, and tryptophan, respectively.[29]

2.2. Optimal Excitation Wavelength of SERS and Thermal Stability of RBDs. The Raman spectra of SARS-CoV-2 RBD on 20 nm AuNPs exhibit strong dependence on the excitation wavelength. Among the three lasers employed, the 785 nm laser provides the best SERS responses for SARS-CoV-2 RBD, while the 633 nm laser offers fewer Raman-active modes in the Raman scattering spectrum in comparison with the 785 nm laser (Figure 2a). In contrast, no Raman peaks are observed under the excitation of the 532 nm laser. Similarly, the RBD of MERS-CoV spike protein also displays an excitation-dependent SERS spectrum when using 20 nm AuNPs as the SERS substrate (Figure 2b). The relative Raman intensities are comparable for 785 and 633 nm laser excitations, whereas the 532 nm excitation still yields no Raman-active modes. The wavelength dependence of SERS is dominated by the plasmon resonance of aggregated AuNPs rather than the optical absorption of RBD samples since both SARS-CoV-2 RBD and MERS-CoV RBD do not have optical absorption at the laser wavelengths (Figure S1).

The absence of the SERS effect for 532 nm excitation can be understood from the electromagnetic mechanism of SERS. For noble metal nanostructures, the Raman enhancement effect is based on the electromagnetic mechanism where the incident photons excite the local surface plasmons of metal nanoparticles generating hot spots at certain locations. The electromagnetic field is enhanced at those hot spots contributing to the Raman enhancement.[31,32] When being excited with a 532 nm laser, the plasmon resonance competes with the interband transition of *d*-band electrons to *sp*-conduction bands.[33,34] This inhibits the scattering of the analyte and leads to a significantly reduced electromagnetic enhancement.[35,36]

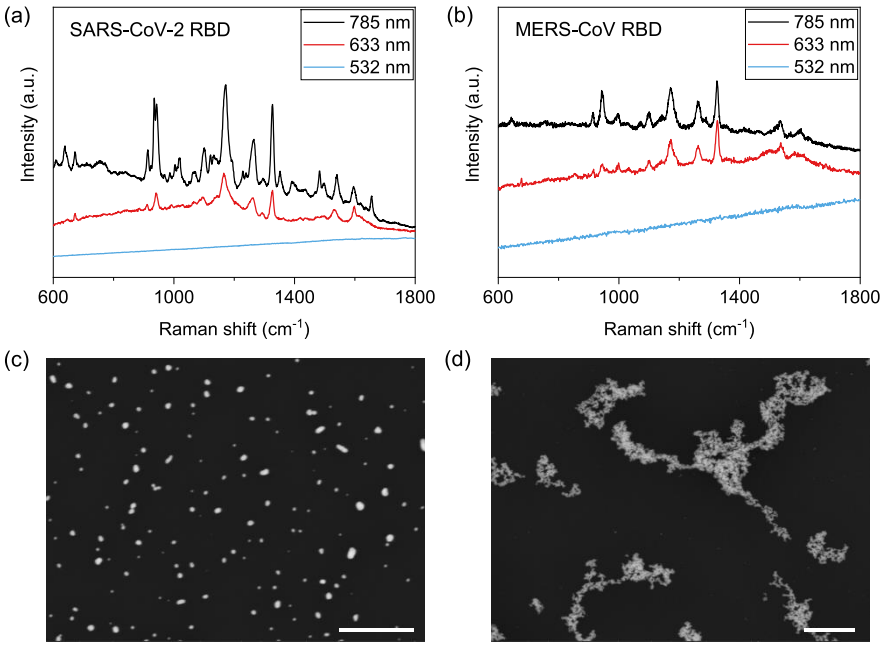


Figure 2. Excitation wavelength dependence of SERS of RBDs. (a) Raman spectra of SARS-CoV-2 RBD and (b) MERS-CoV RBD on AuNPs under 532, 633, and 785 nm lasers. Raman spectra are shifted in the *y* axis. (c) SEM image of backscattered electrons of separated AuNPs on Si before adding RBDs. The scale bar is 500 nm. (d) SEM image of backscattered electrons of AuNP clusters on Si after adding RBDs. The scale bar is 2 μm.

https://doi.org/10.1021/acsphotonics.2c00456
ACS Photonics XXXX, XXX, XXX−XXX

However, the preference for longer excitation wavelengths, as demonstrated by the well-defined Raman peaks under the 633 and 785 nm lasers, can be explained by the red-shift of plasmon resonance. This red-shift of plasmon resonance is revealed by the scattering background in the Raman spectra. As shown in Figure 2a,b, the background under the 633 nm laser rises for higher wavenumbers, while the scattering background under the 785 nm laser decreases for higher wavenumbers, suggesting that the plasmon resonance exists between 633 and 785 nm for the aggregated 20 nm AuNPs. The prevalent factor that needs to be considered for the change of plasmon resonance is the agglomeration of AuNPs, which leads to plasmon resonance at a longer wavelength.[37,38] Particularly for agglomerates composed of more than 15 nanoparticles, it has been demonstrated that the coupling between metal nanoparticles could result in an additional absorption peak at 600−800 nm,[37] closer to the excitation wavelengths of 633 and 785 nm instead of the 532 nm laser. The resonant Raman scattering between the incident photons and the localized surface plasmons of AuNP agglomerates could improve the Raman intensity. The possible influence of AuNP clusters is further supported by the fact that PBS solution is involved in the sample preparation, which promotes the aggregation of nanoparticles.[39] By adding the SARS-CoV-2 RBD in PBS solution to the SERS substrate, Na^+ and K^+ in the PBS solution neutralize the surface charges of AuNPs by binding to the carboxylic acid groups of the citric acid capping, which causes instantaneous and irreversible aggregation of AuNPs.[39] The individual AuNPs on the Si substrate before sample preparation (Figure 2c) and the clustering of AuNPs after adding the RBD (Figure 2d) are visualized by the back-scattered electron images using scanning electron microscopy (SEM).

To ascertain if the wavelength dependence applies to AuNPs of different sizes, we tested the SERS performance of AuNPs with diameters of 60 and 200 nm. The plasmon resonances of unaggregated 60 and 200 nm AuNPs are around 536 and 558 nm, respectively. In contrast to the 20 and 60 nm AuNPs, the absorbance of AuNPs with a diameter of 200 nm extends to 800 nm (Figure S2a,b). As shown in Figure S2c,d, both 60 and 200 nm AuNPs provide stronger Raman enhancement of SARS-CoV-2 RBD under the excitation of a 785 nm laser than 532/633 nm lasers. The level of agglomeration, which is correlated with the mean size of the AuNP aggregates, has been demonstrated to be similar for AuNPs of different diameters.[40] Therefore, we expect that AuNPs of different diameters share the same excitation dependence, which is consistent with our observation. It is worth noting that the aggregation of AuNPs can be effectively manipulated by changing the pH and concentration of NaCl additives. For biological sensing, the participation of biomolecules also affects the electrosteric repulsion and surface adsorption that may ultimately determine the optimal SERS wavelength.[40]

The practicality of biological sensing using SERS relies on the stability of the analyte. Optical excitation oftentimes brings about local heating of the sample that compromises the reliability of the analysis. For Raman scattering, the thermal stability of RBDs is demonstrated by the consistency of Raman spectra under increasing laser power densities. We evaluated different power densities ranging from 0.02 to 0.75 $\mu W/\mu m^2$. In Figure 3a, SARS-CoV-2 RBD shows a consistent Raman intensity profile when being excited with the 785 nm laser and a power density of 0.21 $\mu W/\mu m^2$. However, the Raman

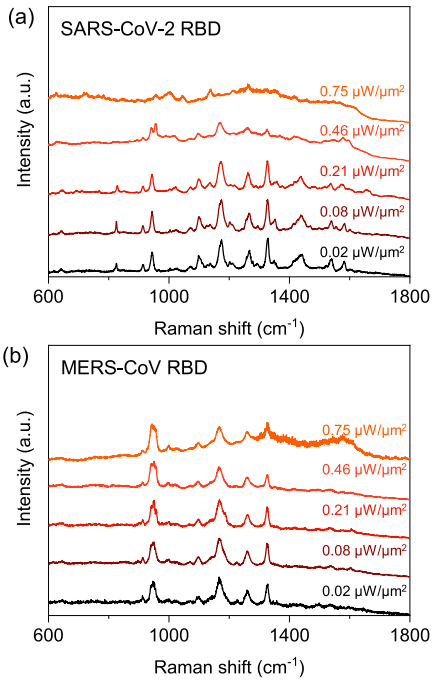


Figure 3. Thermal stability of RBDs. (a) Raman spectra of SARS-CoV-2 RBD and (b) MERS-CoV RBD measured using a 785 nm laser with different power densities. The spectra are normalized and shifted in the *y* axis.

features smear out when the laser power increases to higher than 0.46 $\mu W/\mu m^2$, indicating sample damage due to laser heating. High laser power also results in unwanted fluorescence background. The thermal stability of MERS-CoV RBD is slightly better than SARS-CoV-2 RBD in that the Raman features persist to 0.46 $\mu W/\mu m^2$ and partially to 0.75 $\mu W/\mu m^2$ (Figure 3b). The thermal stability of a protein is correlated with hydrogen bonding and ion pairs.[41] Therefore, the better stability of MERS-CoV RBD suggests that more hydrogen bonding is present in MERS-CoV RBD than in SARS-CoV-2 RBD. To have a comparison of the thermal stability of different kinds of protein peptides, we also examined the Raman spectra of the MERS-CoV S1 mouse monoclonal antibody under different power densities. As shown in Figure S3, the MERS-CoV S1 antibody exhibits Raman peaks at 712, 849, 1003, 1127, 1238, and 1375 cm^{-1}, which are drastically different from the Raman peaks of SARS-CoV-2 RBD and MERS-CoV RBD. MERS-CoV S1 mouse monoclonal antibody is considerably more stable under laser heating as the features in the Raman spectra hold up to 0.92 $\mu W/\mu m^2$. This indicates that the critical power density needs to be evaluated for different kinds of peptides to achieve reliable SERS identification.

2.3. Visualization by Principal Component Analysis and Classification by Machine Learning. The Raman spectrum of a protein is composed of vibrational modes originating from aromatic and nonaromatic amino acid side chains and polypeptide backbones. Thus, it contains structural and compositional information about the nanostructure and is unique to the analyte protein (as shown for 20 amino acids in Figure S4). To elucidate how the distinct Raman features of SARS-CoV-2 RBDs are related to the nanostructure, including chemical composition and structural arrangement, we visualized and analyzed 476 and 324 Raman spectra obtained

https://doi.org/10.1021/acsphotonics.2c00456
ACS Photonics XXXX, XXX, XXX−XXX

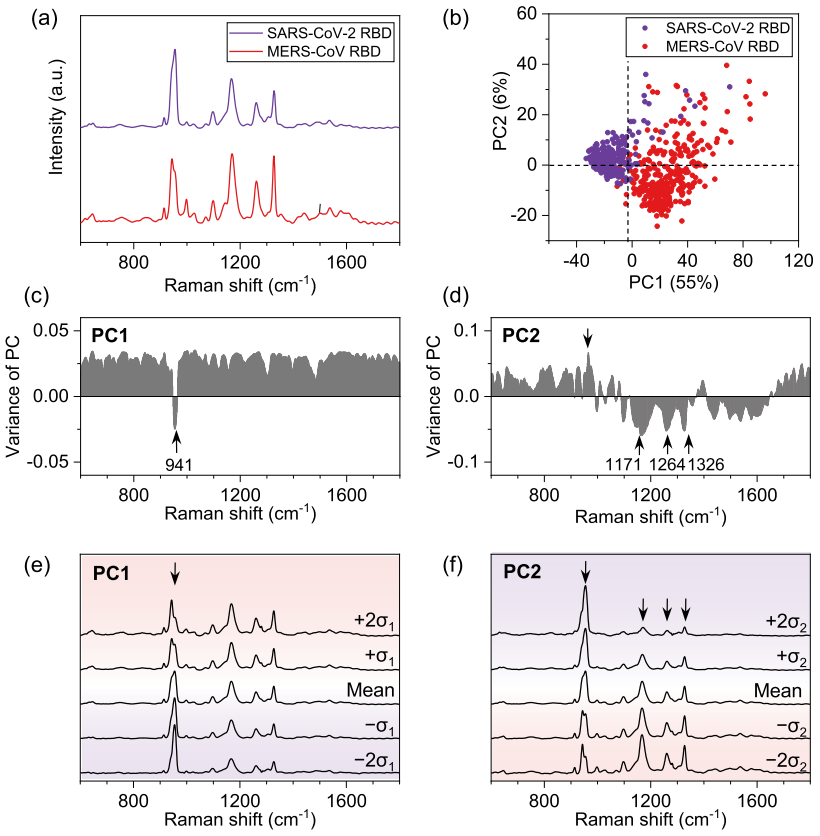


Figure 4. PCA of SARS-CoV-2 RBD and MERS-CoV RBD Raman spectra. (a) The average Raman spectra of SARS-CoV-2 RBD and MERS-CoV RBD. Raman spectra are averaged for 476 preprocessed SARS-CoV-2 RBD spectra and 324 MERS-CoV RBD spectra. (b) Visualization of SARS-CoV-2 RBD and MERS-CoV RBD SERS spectra using PCA. Each point represents a spectrum; SARS-CoV-2 RBD spectra are rendered in purple and MERS-CoV RBD spectra are in red. (c) Variance for each Raman shift along PC1. (d) Variance for each Raman shift along PC2. (e) Reconstructed Raman spectra by varying from the mean spectrum along the PC1 direction. (f) Reconstructed spectra by varying from the mean spectrum along the PC2 direction.

for SARS-CoV-2 RBD and MERS-CoV RBD, respectively, using principal component analysis (PCA), as shown in Figure 4. The average Raman spectra for SARS-CoV-2 and MERS-CoV RBDs are shown in Figure 4a. Each Raman spectrum is represented as a point in a two-dimensional coordinate by PCA in Figure 4b. Principal components 1 (PC1) and 2 (PC2) explain 55 and 6% of the variance among spectra, respectively. As shown in Figure 4b, the Raman spectra of SARS-CoV-2 RBD and MERS-CoV RBD are separable by their values along the axis of PC1: while SARS-CoV-2 RBD has a mean of −17.7 along PC1, MERS-CoV RBD has a mean of 26.1 along PC1. Figure 4c visualizes the variance among all spectra for each Raman shift along PC1. One can see that the variance has a distinctive feature at 941 cm^{-1}, suggesting that the stretching vibration of the α-helix at 941 cm^{-1} significantly contributes to the clustering of Raman spectra by PC1. However, for PC2 in Figure 4d, the dominant characteristics of the variance are located at 1171, 1264, and 1326 cm^{-1} that are assigned to the deformation vibrations of CH_3 (1171 cm^{-1}) and amide III (1264 and 1326 cm^{-1}) (Table 1). The remaining PCs, including PC3 and PC4, mostly describe the spectral noise. Therefore, the Raman spectra of SARS-CoV-2 RBD and MERS-CoV RBD are not well separable along the axes of PC3 and PC4, as shown by the PC1−PC3 and PC1−PC4 plots in Figure S5. We also implemented the partial least squares (PLS) analysis for the Raman spectra of SARS-CoV-2 RBD and MERS-CoV RBD. As shown in Figure S6, the PLS plot of latent variable 1 and latent variable 2 reveals similar patterns to the PCA plot of PC1−PC2 (Figure 4b).

The interpretation of the variance for each PC can be easier to comprehend by reconstructing several spectra that are the result of adding one or two times of the standard deviation σ along the PC direction to the mean spectrum. Details for the reconstruction are described in the Methods section. In Figure 4e, the reconstructed spectra that vary from the mean spectra by $+2\sigma_1$ and $-2\sigma_1$ (as highlighted by red and purple backgrounds, respectively) resemble the Raman spectra of MERS-CoV and SARS-CoV-2 RBDs in the experiments, respectively. This can be easily explained because MERS-CoV RBD spectra have more positive PC1 values compared to SARS-CoV-2 RBD spectra (see Figure 4b). Furthermore, Figure 4e shows that the shape and intensity of the Raman peak at 941 cm^{-1} stemming from the α-helix have obvious changes when varying along the PC1 direction, which is consistent with the variance shown in Figure 4c. The α-helix is the conformation where the hydrogen atoms of the backbone N−H bond to the backbone C═O. It is a common motif in the secondary structure of peptides, which accounts for the stability and robustness to mutation.[42] The difference in the thermal stability of SARS-CoV-2 RBD and MERS-CoV RBD upon laser irradiation could be related to the proximity of the α-helix to the SERS surface. For example, if the α-helix of SARS-CoV-2 RBD is closer to the SERS surface, it could exhibit stronger relative Raman intensity compared to that in

https://doi.org/10.1021/acsphotonics.2c00456
ACS Photonics XXXX, XXX, XXX−XXX

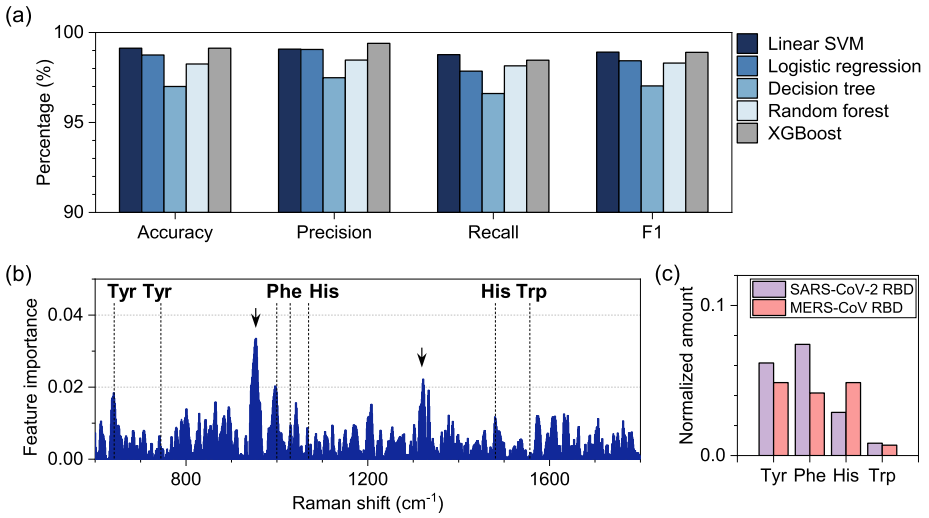


Figure 5. Machine learning classification of SARS-CoV-2 RBD and MERS-CoV RBD. (a) Classification performance of different classifiers. Four metrics including accuracy, precision, recall, and F_1 scores were measured. (b) Feature importance calculated from applying linear SVM to classifying SARS-CoV-2 RBD and MERS-CoV RBD. The dashed lines label the Raman shifts of tyrosine (Tyr), phenylalanine (Phe), histidine (His), and tryptophan (Trp). The arrows label the stretching vibrations of $C^{\alpha}-C^{\beta}$ (α-helix) at 941 cm^{-1} and amide III at 1326 cm^{-1}, respectively. (c) Amounts of tyrosine (Tyr), phenylalanine (Phe), histidine (His), and tryptophan (Trp) are normalized by the total number of amino acids in the RBD based on the sequencing results.

the MERS-CoV RBD. At the same time, this secondary structure is also prone to break down because of the strong interaction with surface plasmons. In a similar way, the reconstructed spectra by varying along the PC2 direction from the mean spectra, as shown in Figure 4f, are consistent with the results of PCA analysis as shown in Figure 4b. SARS-CoV-2 RBD spectra have mostly positive PC2 values (Figure 4b), and thus, the reconstructed spectrum at $+2\sigma_2$ along PC2 (spectrum in the purple background in Figure 4f) resembles the average Raman spectrum of SARS-CoV-2 RBD to a greater extent, whereas MERS-CoV RBD spectra have mostly negative PC2 values (Figure 4b), and thus, the reconstructed spectrum at $-2\sigma_2$ along PC2 (spectrum in the red background in Figure 4f) resembles the average Raman spectrum of MERS-CoV RBD more. The main characteristic differences observable for PC2 are from deformation vibrations of CH_3 at 1171 cm^{-1}, as well as amide III at 1264 and 1326 cm^{-1} that belong to the secondary structure.

In addition, we employed machine learning to classify the Raman spectra of SARS-CoV-2 RBD and MERS-CoV RBD based on their Raman signatures. The full Raman spectra were used as the input for machine learning models, independent of the PCA result, which allows the machine learning models to calculate and provide feature importance values on individual Raman peaks. Five classifiers, including linear support vector machine (SVM), logistic regression, decision tree, random forest, and XGBoost, were used, and their performances were evaluated using four metrics as shown in Figure 5a. All five classifiers achieved accuracy, precision, recall, and F_1 scores higher than 95%. The formulas for the evaluation metrics and information about detailed machine learning parameters are described in the Supporting Information. Among the applied classifiers, linear SVM shows the highest accuracy, recall, and F_1 scores, while XGBoost achieves the highest precision and performs better than decision tree and random forest. Furthermore, the feature weights used by linear SVM for classification enable us to derive useful information from the Raman features of RBDs. The higher feature importance at certain Raman shifts suggests a larger contribution to the classification from the Raman modes located at that Raman shift. The stretching vibrations of $C^{\alpha}-C^{\beta}$ (α-helix) at 941 cm^{-1} and amide III at 1326 cm^{-1} have high feature importance as labeled by the arrows in Figure 5b, in agreement with the PCA analysis (Figure 4c,d). In terms of the Raman modes originating from specific amino acids, the feature importance is low at the Raman shifts of tryptophan (1540 cm^{-1}) due to few amounts of tryptophan in both SARS-CoV-2 RBD and MERS-CoV RBD (Figures 5c and S7). In contrast, the Raman shifts of tyrosine, phenylalanine, and histidine exhibit substantial feature importance corresponding to the different amounts of each amino acid between SARS-CoV-2 RBD and MERS-CoV RBD. Our feature importance analysis based on linear SVM captures the composition of certain amino acids in RBD peptides that provides a better understanding of their roles in the light−matter interaction.

Research endeavors on COVID-19 are urgent, especially when mutations in the genome of SARS-CoV-2 are found at an accelerated pace. There have been new variants of concern that possess amino acid substitutions in their RBDs.[43] For example, the Delta variant harbors defining mutations L452R in the RBD that interacts with ACE2.[44] Some of the amino acid changes are known to partially evade the immune system and some confer stronger binding to the human ACE2 receptor, causing a higher transmission rate.[11] The Omicron variant even possess at least 30 amino acid substitution with 15 of them in the RBD.[12] Under this circumstance, a further refined SERS and machine learning analysis could be developed to analyze the amino acid changes in RBD. For example, the reported amino acid substitutions in the RBD of the Omicron variant spike protein include S375F, N501Y, and Y505H, which represents substitution from serine to phenylalanine at location 375, asparagine to tyrosine at location 501, and tyrosine to histidine at location 505.[12] Those changes in the RBD might lead to spectral modifications at the Raman frequencies of phenylalanine, tyrosine, and histidine, all of which show high importance in the machine learning

https://doi.org/10.1021/acsphotonics.2c00456
ACS Photonics XXXX, XXX, XXX−XXX

classifications (Figure 5b). To extend our knowledge of the purified proteins to viruses, we can apply data correlation analysis to compare the spectra of RBD and viruses so that Raman fingerprints of RBDs that persist in the viruses can be identified. Advanced machine learning models with interpretability[45,46] can be developed to extract Raman features that are invariant to factors unrelated to the virus and extend the analysis to unpurified or even patient samples. The reliability of the machine learning analysis will benefit from viral enrichment using other low-dimensional materials such as carbon nanotube (CNT) arrays. As shown in Figure S8, the overall Raman intensities increase due to the enrichment of local RBD concentration when CNTs are incorporated with AuNPs in the SERS system.[13,47]

3. CONCLUSIONS

In summary, using SERS and machine learning, we obtained a fundamental understanding of the optical properties, thermal stability, and nanostructure of the RBD of the circulating SARS-CoV-2. In particular, we explored the SERS dependence on the excitation wavelength and the thermal stability of SARS-CoV-2 RBD and MERS-CoV RBD. The demonstrated strong dependence on the laser wavelength can be partly explained by the competition between plasmon resonance and interband transition. The 633 and 785 nm lasers were particularly favored because of the red-shifts of plasmon resonance due to nanoparticle aggregation. We also demonstrated that different types of protein including SARS-CoV-2 RBD, MERS-CoV RBD, and MERS-CoV S1 antibody display contrasting stabilities under laser heating. The understanding of the SERS mechanism provides a guideline for the optimal measurement parameters for RBDs using SERS. To clarify the role of RBD nanostructure in the SERS process, we employed PCA and spectrum reconstruction to analyze the Raman fingerprints of two types of RBDs. The major spectral difference in PCA between SARS-CoV-2 RBD and MERS-CoV RBD is attributed to the α-helix, a secondary structure related to the stability and robustness of viruses. Furthermore, machine learning classification offers detailed insights into the secondary structures and amino acids. Using machine learning classifiers, RBDs from SARS-CoV-2 and MERS-CoV can be easily differentiated with accuracy, precision, recall, and F_1 scores all over 95%. The discrimination is attributed to the vibrations of secondary structures including the α-helix at 941 cm^{-1} and amide III at 1326 cm^{-1} and amino acids including tyrosine, phenylalanine, and histidine, which is consistent with sequencing results. Our study offers key insights into the optical properties of SARS-CoV-2 RBD that may facilitate variant surveillance by monitoring the changes in the amino acid composition of the RBD. This analysis approach can be extended to peptides of a wide range of infectious viruses. We anticipate that a further refined SERS and machine learning platform can differentiate SARS-CoV-2 from other viruses rapidly and accurately and can identify new strains of SARS-CoV-2 in the early stage to maintain immunity.

4. METHODS

4.1. SARS-CoV-2 RBD Expression. The expression of the RBD of SARS-CoV-2 spike protein from the WA-1 strain has been recently described.[48] Briefly, a pCAGGS-based plasmid with a recodonized sequence for expressing the signal peptide and the RBD with a fused 6×His tag was obtained from Florian Krammer[49] and was given the unique identifier pSL1510 for this work. Plasmid was transfected using an Expi293 expression system from Thermo Fisher. Transfected Expi293F cells were cultured for 3 days (37 °C, 8% CO_2, in shaker flasks, (120−130 rpm)), at which point the culture supernatant was collected and subjected to metal affinity chromatography (Ni-NTA). Protein was bound to resin in batches at 4 °C for 1 h, extensively washed, and eluted in 57 mM NaH_2PO_4, 30 mM NaCl, and 235 mM Imidazole. Purified SARS-CoV-2 Spike RBD::6×His was dialyzed into 1 × PBS and stored at −80 °C before use. The MERS-CoV RBD in PBS was purchased from SinoBiological, catalog number 40071-V08B1. The MERS-CoV S1 mouse monoclonal antibody was purchased from Invitrogen, catalog number PIMA533278.

4.2. SERS Sample Preparation. The 20 nm gold nanoparticles (AuNPs) were synthesized using the Turkevich method.[50] Then, 50 mL of $HAuCl_4$ solution (0.01% by weight) was first heated to boil, and 0.8 mL of sodium citrate solution (1% by weight) was slowly added to the boiling solution. The mixture was kept boiling for another 5 min and slowly cooled down to room temperature, yielding a bright red solution. The spherical AuNPs with sizes of 60 and 200 nm were purchased from Nanopartz with product numbers A11-60-CIT-DIH-1-25 and A11-200-CIT-DIH-1-25. The nitrogen-doped CNT arrays were synthesized via chemical vapor deposition using benzylamine as the precursor. The details can be found in our previous publications.[13,47] The SERS substrates that contain AuNPs with and without CNTs were prepared by drop-casting the as-synthesized colloidal AuNP solution on top of the CNT arrays and bare Si, respectively, and dried overnight. The SARS-CoV-2 RBD and MERS-CoV RBD in PBS were drop-cast on SERS substrates and dried before Raman measurement.

4.3. Optical Spectroscopy. Raman measurements were performed with a Horiba LabRAM Raman spectrometer. The Raman spectrometer was equipped with a 600/mm grating and a 100× objective with a NA of 0.9. The acquisition time was 30 s. The laser power density was kept at 0.02 $\mu W/\mu m^2$ except for thermal stability measurements to avoid decomposition of the viral protein. For the thermal stability measurement, the laser power density was estimated by $\frac{4P}{\pi d^2}$ in which P is the laser power and d is the effective spot size (diameter of the laser spot). The Raman maps used for machine learning analysis were taken with a step size of 2 μm. The UV−vis spectra of AuNP were recorded in a 10 mm optical path quartz cuvette using a LAMBDA 950 UV/vis/NIR spectrometer in the wavelength range of 400−850 nm. A baseline correction was performed using water.

4.4. Principal Component Analysis. We measured 500 Raman spectra for SARS-CoV-2 RBD and 500 for MERS-CoV RBD for machine learning analysis. In the mapping of Raman spectra, there is no Raman signal from the virus component when the AuNP is not present. Therefore, the blank Raman spectra were first filtered by setting an intensity threshold. Then, we performed normalization on the intensities so that after normalization, the intensities are within the [0, 1] range, having a mean value of zero and a standard deviation of one for the 476 SARS-CoV-2 RBD and 324 MERS-CoV RBD Raman spectra with signals. The standardization changes the visual shape of the spectral yet retains their statistical equivalence. We then applied PCA, a statistical technique to perform data dimensionality reduction, and then visualize the data using two principal components.[51] The first two components explain

https://doi.org/10.1021/acsphotonics.2c00456
ACS Photonics XXXX, XXX, XXX−XXX

approximately 61% of the variance of all spectra. For reconstructing some representative Raman spectra based on PCA, the variance of each principal component was extracted. A representative spectrum was reconstructed by multiplying the variance with its corresponding principal component score and adding that to the mean spectrum. For each principal component, we calculated the standard deviation of its principal component score and showed several reconstructed spectra within two times the standard deviation from the mean spectrum.[52]

4.5. Machine Learning Classification. To classify spectra of different RBDs, we applied five machine learning methods, including linear SVM,[53] logistic regression,[54] decision tree,[55] random forest,[56] and XGBoost,[57] to classify the preprocessed spectra into two classes (the detailed training parameters for each model are included in the Supporting Information). To test the accuracy of our models, we used five-fold cross-validation by splitting the dataset into fivefold and then using each of them for testing while the others for training. The formulas for the machine learning model evaluation metrics including accuracy, precision, recall, and F_1 scores are included in the Supporting Information.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsphotonics.2c00456.

> Ultraviolet−visible spectra of RBD, surface-enhanced Raman scattering (SERS) of SARS-CoV-2 RBD on 60 and 200 nm gold nanoparticles, SERS of MERS-CoV S1 mouse monoclonal antibody, Raman spectra of 20 amino acids, PCA of other components, visualization of SARS-CoV-2 RBD and MERS-CoV RBD spectra using PLS, amino acids of SARS-CoV-2 RBD and MERS-CoV RBD by sequencing, and SERS of SARS-CoV-2 RBD using AuNP on CNT arrays (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Mauricio Terrones − *Department of Chemistry, Department of Physics, Center for 2-Dimensional and Layered Materials, and Department of Materials Science and Engineering, The Pennsylvania State University, University Park, Pennsylvania 16802, United States; Research Initiative for Supra Materials, Shinshu University, Nagano 380-8553, Japan;* orcid.org/0000-0003-0010-2851; Email: mut11@psu.edu

Shengxi Huang − *Department of Electrical Engineering, The Pennsylvania State University, University Park, Pennsylvania 16802, United States; Department of Electrical and Computer Engineering, Rice University, Houston, Texas 77005, United States;* orcid.org/0000-0002-3618-9074; Email: sjh5899@psu.edu, shengxi.huang@rice.edu

Authors

Kunyan Zhang − *Department of Electrical Engineering, The Pennsylvania State University, University Park, Pennsylvania 16802, United States;* orcid.org/0000-0002-6830-409X

Ziyang Wang − *Department of Electrical Engineering, The Pennsylvania State University, University Park, Pennsylvania 16802, United States; Department of Electrical and Computer Engineering, Rice University, Houston, Texas 77005, United States;* orcid.org/0000-0003-1368-4539

He Liu − *Department of Chemistry, The Pennsylvania State University, University Park, Pennsylvania 16802, United States*

Néstor Perea-López − *Department of Physics and Center for 2-Dimensional and Layered Materials, The Pennsylvania State University, University Park, Pennsylvania 16802, United States*

Jeewan C. Ranasinghe − *Department of Electrical Engineering, The Pennsylvania State University, University Park, Pennsylvania 16802, United States; Department of Electrical and Computer Engineering, Rice University, Houston, Texas 77005, United States*

George Bepete − *Department of Chemistry, Department of Physics, and Center for 2-Dimensional and Layered Materials, The Pennsylvania State University, University Park, Pennsylvania 16802, United States;* orcid.org/0000-0002-5562-1125

Allen M. Minns − *Department of Biochemistry and Molecular Biology, Center for Infectious Disease Dynamics and Huck Institutes of the Life Sciences, The Pennsylvania State University, University Park, Pennsylvania 16802, United States*

Randall M. Rossi − *Huck Institutes of the Life Sciences, The Pennsylvania State University, University Park, Pennsylvania 16802, United States*

Scott E. Lindner − *Department of Biochemistry and Molecular Biology, Center for Infectious Disease Dynamics and Huck Institutes of the Life Sciences, The Pennsylvania State University, University Park, Pennsylvania 16802, United States*

Sharon X. Huang − *College of Information Sciences and Technology, The Pennsylvania State University, University Park, Pennsylvania 16802, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/acsphotonics.2c00456

Funding

K.Z. and S.H. acknowledge the support from National Science Foundation Award No. DMR-2011839 through the Penn State MRSEC−Center for Nanoscale Science. We also acknowledge the National Science Foundation EArly-concept Grants for Exploratory Research (EAGER) Grant No. 2030857 and Growing Convergence Research (GCR) Grant No. ECCS-1934977. This work was also partially funded by two Huck/Pennsylvania State University COVID-19 Seed Grants (one to S.E.L. and another to M.T., S.X.H., and S.H.). S.H. acknowledges the support from the National Science Foundation Grant No. ECCS-1943895 and Johnson & Johnson Inc. for the STEM2D Scholar's Award.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We appreciate the contributions of the Sartorius Cell Culture Facility of the Huck Institutes of the Life Sciences at the Pennsylvania State University.

■ REFERENCES

(1) Ortaliza, J.; Orgera, K.; Amin, K.; Cox, C. *COVID-19 Continues To Be a Leading Cause of Death in the U.S. in September 2021.* https://

www.healthsystemtracker.org/brief/covid19-and-other-leading-causes-of-death-in-the-us/ (accessed Nov 18, 2021).
(2) Saad-Roy, C. M.; Morris, S. E.; Metcalf, C. J. E.; Mina, M. J.; Baker, R. E.; Farrar, J.; Holmes, E. C.; Pybus, O. G.; Graham, A. L.; Levin, S. A.; Grenfell, B. T.; Wagner, C. E. Epidemiological and Evolutionary Considerations of SARS-CoV-2 Vaccine Dosing Regimes. *Science* **2021**, *372*, 363−370.
(3) Chen, W.-H.; Strych, U.; Hotez, P. J.; Bottazzi, M. E. The SARS-CoV-2 Vaccine Pipeline: An Overview. *Curr. Trop. Med. Rep.* **2020**, *7*, 61−64.
(4) Dong, Y.; Dai, T.; Wei, Y.; Zhang, L.; Zheng, M.; Zhou, F. A Systematic Review of SARS-CoV-2 Vaccine Candidates. *Signal Transduction Targeted Ther.* **2020**, *5*, 237.
(5) Shang, J.; Ye, G.; Shi, K.; Wan, Y.; Luo, C.; Aihara, H.; Geng, Q.; Auerbach, A.; Li, F. Structural Basis of Receptor Recognition by SARS-CoV-2. *Nature* **2020**, *581*, 221−224.
(6) Gaebler, C.; Wang, Z.; Lorenzi, J. C. C.; Muecksch, F.; Finkin, S.; Tokuyama, M.; Cho, A.; Jankovic, M.; Schaefer-Babajew, D.; Oliveira, T. Y.; Cipolla, M.; Viant, C.; Barnes, C. O.; Bram, Y.; Breton, G.; Hägglöf, T.; Mendoza, P.; Hurley, A.; Turroja, M.; Gordon, K.; et al. Evolution of Antibody Immunity to SARS-CoV-2. *Nature* **2021**, *591*, 639−644.
(7) Wang, P.; Nair, M. S.; Liu, L.; Iketani, S.; Luo, Y.; Guo, Y.; Wang, M.; Yu, J.; Zhang, B.; Kwong, P. D.; Graham, B. S.; Mascola, J. R.; Chang, J. Y.; Yin, M. T.; Sobieszczyk, M.; Kyratsous, C. A.; Shapiro, L.; Sheng, Z.; Huang, Y.; Ho, D. D. Antibody Resistance of SARS-CoV-2 Variants B.1.351 and B.1.1.7. *Nature* **2021**, *593*, 130−135.
(8) Xie, X.; Liu, Y.; Liu, J.; Zhang, X.; Zou, J.; Fontes-Garfias, C. R.; Xia, H.; Swanson, K. A.; Cutler, M.; Cooper, D.; Menachery, V. D.; Weaver, S. C.; Dormitzer, P. R.; Shi, P.-Y. Neutralization of SARS-CoV-2 Spike 69/70 Deletion, E484K and N501Y Variants by BNT162b2 Vaccine-Elicited Sera. *Nat. Med.* **2021**, *27*, 620−621.
(9) Wang, Z.; Schmidt, F.; Weisblum, Y.; Muecksch, F.; Barnes, C. O.; Finkin, S.; Schaefer-Babajew, D.; Cipolla, M.; Gaebler, C.; Lieberman, J. A.; Oliveira, T. Y.; Yang, Z.; Abernathy, M. E.; Huey-Tubman, K. E.; Hurley, A.; Turroja, M.; West, K. A.; Gordon, K.; Millard, K. G.; Ramos, V.; et al. mRNA Vaccine-Elicited Antibodies to SARS-CoV-2 and Circulating Variants. *Nature* **2021**, *592*, 616−622.
(10) Zhou, D.; Dejnirattisai, W.; Supasa, P.; Liu, C.; Mentzer, A. J.; Ginn, H. M.; Zhao, Y.; Duyvesteyn, H. M. E.; Tuekprakhon, A.; Nutalai, R.; Wang, B.; Paesen, G. C.; Lopez-Camacho, C.; Slon-Campos, J.; Hallis, B.; Coombes, N.; Bewley, K.; Charlton, S.; Walter, T. S.; Skelly, D.; et al. Evidence of Escape of SARS-CoV-2 Variant B.1.351 from Natural and Vaccine-Induced Sera. *Cell* **2021**, *184*, 2348−2361.e6.
(11) Starr, T. N.; Greaney, A. J.; Dingens, A. S.; Bloom, J. D. Complete Map of SARS-CoV-2 RBD Mutations That Escape the Monoclonal Antibody LY-CoV555 and Its Cocktail with LY-CoV016. *Cell Rep. Med.* **2021**, *2*, No. 100255.
(12) Centers for Disease Control and Prevention. *Science Brief: Omicron (B.1.1.529) Variant*. https://www.cdc.gov/coronavirus/2019-ncov/science/science-briefs/scientific-brief-omicron-variant.html (accessed Dec 6, 2021).
(13) Yeh, Y.-T.; Gulino, K.; Zhang, Y.; Sabestien, A.; Chou, T.-W.; Zhou, B.; Lin, Z.; Albert, I.; Lu, H.; Swaminathan, V.; Ghedin, E.; Terrones, M. A Rapid and Label-Free Platform for Virus Capture and Identification from Clinical Samples. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117*, 895.
(14) Huang, S.; Pandey, R.; Barman, I.; Kong, J.; Dresselhaus, M. Raman Enhancement of Blood Constituent Proteins Using Graphene. *ACS Photonics* **2018**, *5*, 2978−2982.
(15) Zong, C.; Premasiri, R.; Lin, H.; Huang, Y.; Zhang, C.; Yang, C.; Ren, B.; Ziegler, L. D.; Cheng, J.-X. Plasmon-Enhanced Stimulated Raman Scattering Microscopy with Single-Molecule Detection Sensitivity. *Nat. Commun.* **2019**, *10*, 5318.
(16) Nie, S.; Emory, S. R. Probing Single Molecules and Single Nanoparticles by Surface-Enhanced Raman Scattering. *Science* **1997**, *275*, 1102.
(17) Zhang, X.; Zheng, Y.; Liu, X.; Lu, W.; Dai, J.; Lei, D. Y.; MacFarlane, D. R. Hierarchical Porous Plasmonic Metamaterials for Reproducible Ultrasensitive Surface-Enhanced Raman Spectroscopy. *Adv. Mater.* **2015**, *27*, 1090−1096.
(18) Silver, A.; Kitadai, H.; Liu, H.; Granzier-Nakajima, T.; Terrones, M.; Ling, X.; Huang, S. Chemical and Bio Sensing Using Graphene-Enhanced Raman Spectroscopy. *Nanomaterials* **2019**, *9*, 516.
(19) Xu, W.; Ling, X.; Xiao, J.; Dresselhaus, M. S.; Kong, J.; Xu, H.; Liu, Z.; Zhang, J. Surface Enhanced Raman Spectroscopy on a Flat Graphene Surface. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 9281−9286.
(20) Ling, X.; Xie, L.; Fang, Y.; Xu, H.; Zhang, H.; Kong, J.; Dresselhaus, M. S.; Zhang, J.; Liu, Z. Can Graphene Be Used as a Substrate for Raman Enhancement? *Nano Lett.* **2010**, *10*, 553−561.
(21) Huang, S.; Ling, X.; Liang, L.; Song, Y.; Fang, W.; Zhang, J.; Kong, J.; Meunier, V.; Dresselhaus, M. S. Molecular Selectivity of Graphene-Enhanced Raman Scattering. *Nano Lett.* **2015**, *15*, 2892−2901.
(22) Ye, J.; Yeh, Y.-T.; Xue, Y.; Wang, Z.; Zhang, N.; Liu, H.; Zhang, K.; Ricker, R.; Yu, Z.; Roder, A.; Perea Lopez, N.; Organtini, L.; Greene, W.; Hafenstein, S.; Lu, H.; Ghedin, E.; Terrones, M.; Huang, S.; Huang, S. X. Accurate Virus Identification with Interpretable Raman Signatures by Machine Learning. *Proc. Natl. Acad. Sci. U. S. A.* **2022**, *119*, No. e2118836119.
(23) Zhang, D.; Zhang, X.; Ma, R.; Deng, S.; Wang, X.; Wang, X.; Zhang, X.; Huang, X.; Liu, Y.; Li, G.; Qu, J.; Zhu, Y.; Li, J. Ultra-Fast and Onsite Interrogation of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) in Waters via Surface Enhanced Raman Scattering (SERS). *Water Res.* **2021**, *200*, No. 117243.
(24) 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. COVID-19 Salivary Raman Fingerprint: Innovative Approach for the Detection of Current and Past SARS-CoV-2 Infections. *Sci. Rep.* **2021**, *11*, 4943.
(25) 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.; Xue, P.; et al. On-Site Detection of SARS-CoV-2 Antigen by Deep Learning-Based Surface-Enhanced Raman Spectroscopy and Its Biochemical Foundations. *Anal. Chem.* **2021**, *93*, 9174−9182.
(26) Tai, W.; Zhang, X.; Drelich, A.; Shi, J.; Hsu, J. C.; Luchsinger, L.; Hillyer, C. D.; Tseng, C.-T. K.; Jiang, S.; Du, L. A Novel Receptor-Binding Domain (RBD)-Based mRNA Vaccine against SARS-CoV-2. *Cell Res.* **2020**, *30*, 932−935.
(27) Lederer, K.; Castaño, D.; Gómez Atria, D.; Oguin, T. H.; Wang, S.; Manzoni, T. B.; Muramatsu, H.; Hogan, M. J.; Amanat, F.; Cherubin, P.; Lundgreen, K. A.; Tam, Y. K.; Fan, S. H. Y.; Eisenlohr, L. C.; Maillard, I.; Weissman, D.; Bates, P.; Krammer, F.; Sempowski, G. D.; Pardi, N.; et al. SARS-CoV-2 mRNA Vaccines Foster Potent Antigen-Specific Germinal Center Responses Associated with Neutralizing Antibody Generation. *Immunity* **2020**, *53*, 1281−1295.e5.
(28) Awada, C.; Abdullah, M. M.; Traboulsi, H.; Dab, C.; Alshoaibi, A. SARS-CoV-2 Receptor Binding Domain as a Stable-Potential Target for SARS-CoV-2 Detection by Surface-Enhanced Raman Spectroscopy. *Sensors* **2021**, *21*, 4617.
(29) 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. Detection of SARS-CoV-2 and Its S and N Proteins Using Surface Enhanced Raman Spectroscopy. *RSC Adv.* **2021**, *11*, 25788−25794.
(30) Kurouski, D.; Postiglione, T.; Deckert-Gaudig, T.; Deckert, V.; Lednev, I. K. Amide I Vibrational Mode Suppression in Surface (SERS) and Tip (TERS) Enhanced Raman Spectra of Protein Specimens. *Analyst* **2013**, *138*, 1665−1673.

https://doi.org/10.1021/acsphotonics.2c00456
ACS Photonics XXXX, XXX, XXX−XXX

(31) Albrecht, M. G.; Creighton, J. A. Anomalously Intense Raman Spectra of Pyridine at a Silver Electrode. *J. Am. Chem. Soc.* **1977**, *99*, 5215−5217.
(32) Philpott, M. R. Effect of Surface Plasmons on Transitions in Molecules. *J. Chem. Phys.* **1975**, *62*, 1812−1817.
(33) Amendola, V.; Pilot, R.; Frasconi, M.; Maragò, O. M.; Iatì, M. A. Surface Plasmon Resonance in Gold Nanoparticles: A Review. *J. Phys.: Condens. Matter* **2017**, *29*, 203002.
(34) Link, S.; El-Sayed, M. A. Optical Properties and Ultrafast Dynamics of Metallic Nanocrystals. *Annu. Rev. Phys. Chem.* **2003**, *54*, 331−366.
(35) Schreiber, B.; Gkogkou, D.; Dedelaite, L.; Kerbusch, J.; Hübner, R.; Sheremet, E.; Zahn, D. R. T.; Ramanavicius, A.; Facsko, S.; Rodriguez, R. D. Large-Scale Self-Organized Gold Nanostructures with Bidirectional Plasmon Resonances for SERS. *RSC Adv.* **2018**, *8*, 22569−22576.
(36) Baia, M.; Toderas, F.; Baia, L.; Popp, J.; Astilean, S. Probing the Enhancement Mechanisms of SERS with *P*-Aminothiophenol Molecules Adsorbed on Self-Assembled Gold Colloidal Nanoparticles. *Chem. Phys. Lett.* **2006**, *422*, 127−132.
(37) Zook, J. M.; Rastogi, V.; MacCuspie, R. I.; Keene, A. M.; Fagan, J. Measuring Agglomerate Size Distribution and Dependence of Localized Surface Plasmon Resonance Absorbance on Gold Nanoparticle Agglomerate Size Using Analytical Ultracentrifugation. *ACS Nano* **2011**, *5*, 8070−8079.
(38) Kim, T.; Lee, K.; Gong, M.-S.; Joo, S.-W. Control of Gold Nanoparticle Aggregates by Manipulation of Interparticle Interaction. *Langmuir* **2005**, *21*, 9524−9528.
(39) Albanese, A.; Chan, W. C. W. Effect of Gold Nanoparticle Aggregation on Cell Uptake and Toxicity. *ACS Nano* **2011**, *5*, 5478−5489.
(40) Bélteky, P.; Rónavári, A.; Zakupszky, D.; Boka, E.; Igaz, N.; Szerencsés, B.; Pfeiffer, I.; Vágvölgyi, C.; Kiricsi, M.; Kónya, Z. Are Smaller Nanoparticles Always Better? Understanding the Biological Effect of Size-Dependent Silver Nanoparticle Aggregation under Biorelevant Conditions. *Int. J. Nanomed.* **2021**, *16*, 3021−3040.
(41) Vogt, G.; Woell, S.; Argos, P. Protein Thermal Stability, Hydrogen Bonds, and Ion Pairs. *J. Mol. Biol.* **1997**, *269*, 631−643.
(42) Rocklin Gabriel, J.; Chidyausiku Tamuka, M.; Goreshnik, I.; Ford, A.; Houliston, S.; Lemak, A.; Carter, L.; Ravichandran, R.; Mulligan Vikram, K.; Chevalier, A.; Arrowsmith Cheryl, H.; Baker, D. Global Analysis of Protein Folding Using Massively Parallel Design, Synthesis, and Testing. *Science* **2017**, *357*, 168−175.
(43) World Health Organization. *Tracking SARS-CoV-2 variants*. https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/ (accessed Dec 6, 2021).
(44) Hoffmann, M.; Hofmann-Winkler, H.; Krüger, N.; Kempf, A.; Nehlmeier, I.; Graichen, L.; Arora, P.; Sidarovich, A.; Moldenhauer, A.-S.; Winkler, M. S.; Schulz, S.; Jäck, H.-M.; Stankov, M. V.; Behrens, G. M. N.; Pöhlmann, S. SARS-CoV-2 Variant B.1.617 Is Resistant to Bamlanivimab and Evades Antibodies Induced by Infection and Vaccination. *Cell Rep.* **2021**, No. 109415.
(45) Zolna, K.; Geras, K. J.; Cho, K. Classifier-Agnostic Saliency Map Extraction. *Comput. Vis. Image Underst.* **2020**, *196*, No. 102969.
(46) Selvaraju, R. R.; Cogswell, M.; Das, A.; Vedantam, R.; Parikh, D.; Batra, D. In *Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization*, 2017 IEEE International Conference on Computer Vision (ICCV), 22−29 Oct. 2017, 2017; pp 618−626.
(47) Yeh, Y.-T.; Tang, Y.; Sebastian, A.; Dasgupta, A.; Perea-Lopez, N.; Albert, I.; Lu, H.; Terrones, M.; Zheng, S.-Y. Tunable and Label-Free Virus Enrichment for Ultrasensitive Virus Detection Using Carbon Nanotube Arrays. *Sci. Adv.* **2016**, *2*, No. e1601026.
(48) Gontu, A.; Srinivasan, S.; Salazar, E.; Nair, M. S.; Nissly, R. H.; Greenawalt, D.; Bird, I. M.; Herzog, C. M.; Ferrari, M. J.; Poojary, I.; Katani, R.; Lindner, S. E.; Minns, A. M.; Rossi, R.; Christensen, P. A.; Castillo, B.; Chen, J.; Eagar, T. N.; Yi, X.; Zhao, P.; et al. Limited Window for Donation of Convalescent Plasma with High Live-Virus Neutralizing Antibody Titers for COVID-19 Immunotherapy. *Commun. Biol.* **2021**, *4*, 267.
(49) Amanat, F.; Stadlbauer, D.; Strohmeier, S.; Nguyen, T. H. O.; Chromikova, V.; McMahon, M.; Jiang, K.; Arunkumar, G. A.; Jurczyszak, D.; Polanco, J.; Bermudez-Gonzalez, M.; Kleiner, G.; Aydillo, T.; Miorin, L.; Fierer, D. S.; Lugo, L. A.; Kojic, E. M.; Stoever, J.; Liu, S. T. H.; Cunningham-Rundles, C.; et al. A Serological Assay to Detect SARS-CoV-2 Seroconversion in Humans. *Nat. Med.* **2020**, *26*, 1033−1036.
(50) Frens, G. Controlled Nucleation for the Regulation of the Particle Size in Monodisperse Gold Suspensions. *Nat. Phys. Sci.* **1973**, *241*, 20−22.
(51) Schölkopf, B.; Smola, A. J., *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. The MIT Press, 2018.
(52) Cootes, T. F.; Taylor, C. J.; Cooper, D. H.; Graham, J. Active Shape Models-Their Training and Application. *Comput. Vis. Image Underst.* **1995**, *61*, 38−59.
(53) Zhang, Y., Support Vector Machine Classification Algorithm and Its Application. In *ICICA 2012. Communications in Computer and Information Science*, Springer: 2012; pp 179−186.
(54) Friedman, J.; Hastie, T.; Tibshirani, R. Additive Logistic Regression: A Statistical View of Boosting. *Ann. Stat.* **2000**, *28*, 337−407.
(55) Dietterich, T. G., Ensemble Methods in Machine Learning. In *Multiple Classifier Systems*, Springer: 2000; pp 1−15.
(56) Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5−32.
(57) Chen, T.; Guestrin, C., XGBoost: A Scalable Tree Boosting System. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Association for Computing Machinery: San Francisco, California, USA, 2016; pp 785−794.

https://doi.org/10.1021/acsphotonics.2c00456
ACS Photonics XXXX, XXX, XXX−XXX