Characterization of Ras Protein Using Raman Spectroscopy

By

Makunda Aryal

July 2024

Director of Dissertation: Yong-Qing Li, PhD

Major Department: Physics

ABSTRACT

Ras is a family of related proteins expressed in all animal cells, which is small GTPase anchors to the plasma membrane where it regulates signal transduction for cell proliferation, survival, and other cell functions. Mutation of Ras makes it lost the ability to regulate signal transduction and causes cancer. Three different types of Ras (KRAS, HRAS, NRAS) and their isoforms have been identified in different human cancers. Understanding Ras protein functioning mechanisms is crucial in developing cancer drug and cancer therapy at the molecular level. At present, X-ray crystallography and NMR are the two major techniques to effectively measure Ras protein structures and its conformation, but they cannot measure Ras protein conformations in cells under physiological conditions. In this study, we intend to develop label-free Raman spectroscopy methods to characterize and quantitate Ras isoforms and their structural conformations with their unique finger-printing vibrational spectra in vivo and in vitro, as well as to monitor the small drug molecule's binding with Ras molecule in live cells, which is new to Ras based cancer biology.

The first part of this dissertation work is to explore the capability of near-infrared Raman spectroscopy and multivariate analysis techniques for the characterization and differentiation of Ras isoforms and their structural conformations with their finger-printing Raman spectra. We have collected Raman spectra from different Ras isoforms. We have developed a principal component analysis (PCA), a discriminant analysis of principal component (DAPC), and Raman barcode methods, and demonstrated that Ras isoforms can be differentiated by their Raman spectra. We have studied the structural conformations of Ras proteins with GDP and GTP loading. By deconvolution of Raman spectra, we were able to obtain new information about Ras conformational structures including protein's secondary structures, hydrogen bonding condition of phenol side chain, and hydrophobic nature of tyrosine doublet for each Ras isoforms. We have applied Raman spectroscopy for the study of the specific inhibitor (ARS1620 drug) and KRAS G12C interactions, offering new insights into Ras-targeted cancer therapies. These data allow generating a library of Raman spectra of Ras isoforms in vitro, which may serve as the control and platform for developing methods to detect the Ras Raman fingerprint inside a cell.

The second part is to develop a vacuum-enhanced micro-Raman spectroscopy (VERS) for the detection of analyte molecules at relatively low concentrations in an aqueous medium. Given the low concentration of Ras proteins in vitro and in vivo, we aim to develop a novel label-free enhanced Raman spectroscopy for detecting proteins and other molecules at low concentrations. The VERS technique relies on the increase in the molecule's concentration within the micron-sized excitation volume of laser focus by vacuum evaporation of solvents and does not cause the alteration in normal Raman spectra. We demonstrate that an enhancement factor of ~10^3-10^4 was observed for glucose and protein samples, and the enhancement factor depends on the size of the sample holder and the volume of the liquid sample used. We show that VERS can be used to detect

ciprofloxacin antibiotics in human urine at a level of 2 μg/ml. We also show that the VERS technique is particularly useful for detecting biomolecules resolved in a solvent such as dimethyl sulfoxide (DMSO) that has an intense Raman background due to the evaporation of the solvent.

The third part is to develop the surface-enhanced Raman scattering (SERS) spectroscopy and SERS-based super-resolution Raman imaging techniques for the detection and chemical imaging of analytes at the single-molecule level. Due to the low probability of intrinsic Raman scattering and the nanomolar intracellular concentration of Ras protein, we introduce nano-sized metallic particles to bind with the target molecules, using their SERS signals for ultrahigh sensitivity detection and precise spatial location of molecules. We built a home-made experimental setup that allows simultaneous observation of the Raman images of hot spots and acquisition of time-lapse Raman spectra of a single hot particle. We explored the dynamic behavior of SERS of hot particles at the single-molecule level, focusing on the influence of environmental conditions and laser power on the stability of SERS Raman signals. It was observed that SERS intensity fluctuations were more pronounced at the atmospheric pressure than at a low pressure. The experimental data further revealed that as the laser power increased (~ 8.9 μW and above), there was a notable increase in SERS signal fluctuations and blinking, probably due to thermal or photo-induced changes in nanoparticle clusters. We proved that the use of bright field microscopy and Raman imaging for tracking hotspots together with ThunderSTORM software, was effective in accurately identifying and analyzing the change in hot particle's center position. These findings underscore the importance of precise experimental control on both laser parameters and environmental conditions to optimize SERS at single molecule levels, which advances our understanding of SERS at the nanoscale, promoting its application in nanotechnology and materials science.

Characterization of Ras Protein Using Raman Spectroscopy

A Dissertation

Presented to the Faculty of the Department of Physics

East Carolina University

In Partial Fulfillment of the Requirements for the Degree

Doctor of Philosophy in Biomedical Physics

By

Makunda Aryal

July, 2024

Director of Dissertation: Yong-Qing Li, PhD

Dissertation Committee Members:

Michael Dingfelder, PhD

Nathan E. Hudson, PhD

Juan Beltran-Huarac, PhD

Shouquan Huo, PhD

ACKNOWLEDGMENTS

First and foremost, I extend my deepest gratitude to my supervisor, Dr. Yong Qing Li, for his invaluable guidance, diligent supervision, and unwavering support throughout my research journey. His expertise and insights have been instrumental in shaping both my academic endeavors and this dissertation.

I am immensely thankful to the distinguished members of my dissertation committee: Dr. Michael Dingfelder, Dr. Nathan E. Hudson, Dr. Juan Beltran-Huarac, and Dr. Shouquan Huo. Their expertise and thoughtful feedback have greatly contributed to the refinement of my research.

My sincere appreciation goes to the physics technical staff, Mr. Chris Bonnerup, Mr. Gene Oakley, and Mr. Michael Pawluch, whose technical expertise and assistance were crucial in navigating the complex experimental phases of my study. I am also grateful to all the faculty and administrative staff of the Physics Department at ECU for their support and encouragement throughout my academic journey.

Lastly, I wish to express my heartfelt thanks to my family. Their endless love, encouragement, and sacrifice have been the bedrock of my perseverance and success. Without their support, reaching this stage in my academic career would have been unimaginable.

Table of Contents

List of Tables

List of Figures

List of Symbols and Abbreviations

AI	Artificial intelligence
Amide	Organic functional group in protein
BSA	Bovine serum albumin
CCD	Charge coupled device
DAPC	Discrimination analysis of principal component
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
EF	Enhancement factor
EM	Electromagnetic
EMCCD	Electron multiplying charge coupled device
FERS	Fiber enhancement Raman spectroscopy
GAP	GTPase activating protein
GDP	Guanosine diphosphate
GEF	Guanine nucleotide exchange factor
GppNHp	Guanosine-5'-[(β,γ)-imido]triphosphate (a non-hydrolyzable GTP analog)
GTP	Guanosine triphosphate
HOMO	Highest occupied molecular orbital
HRAS	Harvey rat sarcoma viral oncogene homolog
KRAS	Kirsten rat sarcoma viral oncogene homolog
LDA	Linear discriminant analysis
LSRP	Localized surface plasmon resonance
LUMO	Lowest unoccupied molecular orbital
NP	Nano particle
NRAS	Neuroblastoma RAS viral oncogene homolog
PCA	Principal component analysis
R6G	Rhodamine 6G
RAS	Rat sarcoma
RB	Rhodamine B

RNA	Ribonucleic acid
RR	Resonance Raman
RTK	Receptor tyrosine kinase
SERS	Surface enhancement Raman spectroscopy
TERS	Tip enhancement Raman spectroscopy
Trp	Tryptophan
Tyr	Tyrosine
UV VIZ	Ultraviolet visible (spectroscopy)
VERS	Vacuum enhancement Raman spectroscopy

Chapter 1. Introduction and Theory

1.1 Introduction

Ras is a family of related proteins expressed in all animal cells and organs. The Ras superfamily belongs to a class of proteins called small GTPases in all eukaryotes and are involved in transmitting signals within cells [1-3]. Ras superfamily possesses 30-55 % sequence homology among its members [4]. The Ras gene structure contains a G-domain and carboxy-terminal. The G-domain corresponds to the initial 166-168 residues, which include the phosphate-binding loop (P-loop, residues 10-17), switch I (residues 30-38), switch II (residues 60-76), and the base binding loops (residues 116-120 and 145-147) [5]. The carboxy-terminal contains a hypervariable region (HVR), including CAAX, where C is a cysteine amino acid, A is an aliphatic amino acid, and X is an amino acid [6]. There are mainly three Ras isoforms in human cells: K-Ras, H-Ras, and N-Ras. They regulate the signal transduction of cell proliferation, survival, and differentiation [2]. These three proteins share high homology except for the C-terminus hypervariable region, which performs the specific function of each protein. These Ras proteins cycle between the GTP (active state) and GDP (inactive state) bound states and act as a binary molecular switch that functions in a signaling event [2]. Any mutation in the Ras gene structure results in a permanently active state, leading to uncontrolled cell growth and finally forming a tumor [7].

RAS mutations occur in approximately 19 % of all cancers. KRAS is the most frequently mutated isoform, then NRAS and HRAS [8,9]. RAS isoform mutations show selectivity in various human cancers. For example, pancreatic, lung, and colon cancers show the KRAS mutations with nearly 66.1 %, 16.5 %, and 30.3%, respectively. NRAS mutation occurs in chronic myelomonocytic leukemia (CMML) and acute myeloid leukemia (AML) with 13.1 % and 13.6%

respectively. The malignant melanoma, thyroid carcinoma, and larynx carcinoma have NRAS mutation with frequencies of 18.6 %, 8.1 %, and 9.7%, respectively [10]. The study has shown that approximately 80 % of mutations reside at the G12 of the KRAS [11].

Raman spectroscopy is a spectroscopic technique typically used to determine vibrational modes of molecules, which is commonly used to provide a structural fingerprint by which molecules can be identified [12]. After discovering the Raman effect by C.V. Raman in 1928, many improvements and applications have been developed in Raman spectroscopy to detect and characterize molecules. Since the normal Raman scattering is a very weak phenomenon, we need to enhance the Raman signal to get useful information. Many techniques have been developed [13] including surface-enhanced Raman spectroscopy (SERS) [14, 15], tip enhancement Raman spectroscopy (TERS) [16, 17], fiber enhancement Raman spectroscopy (FERS) [18, 19], UV resonance Raman spectroscopy [20, 21] etc. Each technique has its advantages and limitations.

Surface-enhanced Raman spectroscopy (SERS) employs metal nanoparticles to boost the Raman signals of analyte molecules, making it possible to detect analytes at the single-molecule level [22]. This capability has made SERS a prominent technique in ultra-sensitive chemical detection. Additionally, Raman imaging, which is also based on the Raman effect, allows for the visualization of specific molecules within a sample. At the single-molecule scale, the SERS signals are known to exhibit fluctuations, the causes of which remain a mystery, sparking ongoing research and interest in the scientific community.

Since the different Ras proteins are responsible for the different types of cancer, it is crucial to rapidly differentiate the different Ras and its isoforms. Raman spectroscopy and multivariant analysis techniques (including PCA and DAPC, etc.) have great potential to differentiate Ras isoforms. Also, characterization of Ras proteins between the GTP (active state) and GDP (inactive

state) bound states is potentially possible via Raman spectroscopy. In addition, it is important to find the structural information (like secondary structure) that provides the folding and unfolding state of proteins. This could be estimated by the peak fitting of the amide I band of the Raman spectrum of Ras. Raman spectroscopy can also provide information on hydrogen bonding conditions and hydrophobicity of the protein environment, which helps to study its interaction with other molecules existing in the environment.

Normally, the concentration of Ras protein is found in low concentration level and the normal Raman is a weak phenomenon. That is why we need to develop a technique to enhance the Ras protein at low concentrations liquid state. Also, the concentration of Ras protein in human cell is in the range of 33nM – 500nM [23]. To detect the Ras at even nM level we require a highly enhanced technique such as SERS.

1.2 Theory

1.2.1 Raman Spectroscopy

The phenomenon of Raman scattering was first observed experimentally by the physicists C.V Raman and K.S. Krishnan in 1928 [24]. It is based on the inelastic scattering of incident radiation by its interaction with vibrating molecules. Raman spectroscopy measures the scattering of light by matter. When the laser interacts with the molecules, the molecules scatter the incident photons. There are two different types of scattering: elastic and inelastic scattering. In the elastic scattering, also called Rayleigh scattering, the scattered photon has the same frequency as the incident photon. In the inelastic scattering, the emitted photon has a different frequency from the incident photon, and this type of scattering is called Raman scattering.

The inelastic scattering or Raman scattering, happens through two processes: 1) Stokes process, if the frequency of the emitted radiation is lower than that of the incident photon; 2) Anti-Stokes scattering, if the frequency of the emitted radiation is higher than that of the incident photon.

- **Classical theory of Raman scattering**

The classical theory of Raman scattering can be explained by molecular polarizability of vibrating molecules. The positively charged nucleus and negatively charged electrons are displaced when the molecule is placed in an electric field. As a result of the separation of the charged particles, an electric dipole moment is induced in the molecule. If *E* is the electric field strength and μ is the induced dipole moment, then

$$\mu = \alpha E \tag{1}$$

where α is the molecular polarizability of the molecule. The electric field of an electromagnetic wave of frequency ν experienced by each molecule to is given by

$$E = E_0 \cos(2\pi\nu t) \tag{2}$$

where $\boldsymbol{E}_0$ is the amplitude of the electromagnetic wave. Substituting Equation (2) in (1) leads to

$$\mu = \alpha E_0 \cos(2\pi\nu t) \tag{3}$$

From Eq. (3), we can see that the time-dependent molecular dipole moment emits the radiation at the same frequency, which is the classical explanation of the Rayleigh scattering. For vibrating molecules, the relative location of each atom is time-dependent such that its molecular polarizability α is time-dependent as well. The ability to perturb the local electron cloud depends on the relative location of each atom, and hence the polarizability is the function of the instantaneous position of the atoms. Hence the polarizability changes with small displacement from the equilibrium position, which is given by

$$\alpha = \alpha_0 + (q - q_{eq})\frac{\partial \alpha}{\partial q} \quad (4)$$

Here, α_0 is the equilibrium polarizability, and q_{eq} and q are bond lengths at equilibrium position and any instant, respectively. If the molecule executes a simple harmonic motion, then the displacement can be written as

$$q - q_0 = q_{max}\cos(2\pi\nu_{vib}t) \quad (5)$$

where ν_{vib} is the vibrational frequency of a molecule and q_{max} is the maximum separation distance of atoms from the equilibrium position. Substituting Eq. (5) in Eq. (4), we get

$$\alpha = \alpha_0 + \frac{\partial \alpha}{\partial q} q_{max}\cos(2\pi\nu_{vib}t) \quad (6)$$

substituting Eq.(6) in (3), we get

$$\mu = E_0\cos(2\pi\nu t)\left[\alpha_0 + \frac{\partial \alpha}{\partial q} q_{max}\cos(2\pi\nu_{vib}t)\right] \quad (7)$$

$$\mu = \alpha_0 E_0\cos(2\pi\nu t) + E_0\left(\frac{\partial \alpha}{\partial q}\right)q_{max}\cos(2\pi\nu t)\cos(2\pi\nu_{vib}t)] \quad (8)$$

$$\mu = \alpha_0 E_0\cos(2\pi\nu t) + \frac{E_0}{2}q_{max}\left(\frac{\partial \alpha}{\partial q}\right)\cos[2\pi(\nu - \nu_{vib})t] \quad (9)$$

$$+ \frac{E_0}{2} q_{max}\left(\frac{\partial \alpha}{\partial q}\right)\cos[2\pi(\nu + \nu_{vib})t]$$

In Eq. (9), the first term represents the Raleigh scattering. The second and third terms represent Stokes and anti-Stokes scattering, respectively, which are the inelastic scattering and are changed by the vibrotational frequency of the molecular bond.

- **Quantum theory of Raman scattering**

According to the quantum theory, the molecules exist in the quantized energy levels corresponding to all possible stationary states of the molecule. When the incident radiation with

a photon energy hν is an incident on the sample, and the collision occurs between the incident photons and the molecules. Then the molecules in the ground state are excited to a virtual state. Since the lifetime of the virtual state is very small, the excited molecules return to a vibrational level of the ground state. During this collision, if the molecule's energy remains unchanged after the interaction with the photon, this causes the Rayleigh scattering. Most of the time, the molecules undergo elastic scattering. In very rarer cases, few molecules undergo inelastic scattering, in which there is a transfer of energy between molecules and scattered photons. When the molecule gains

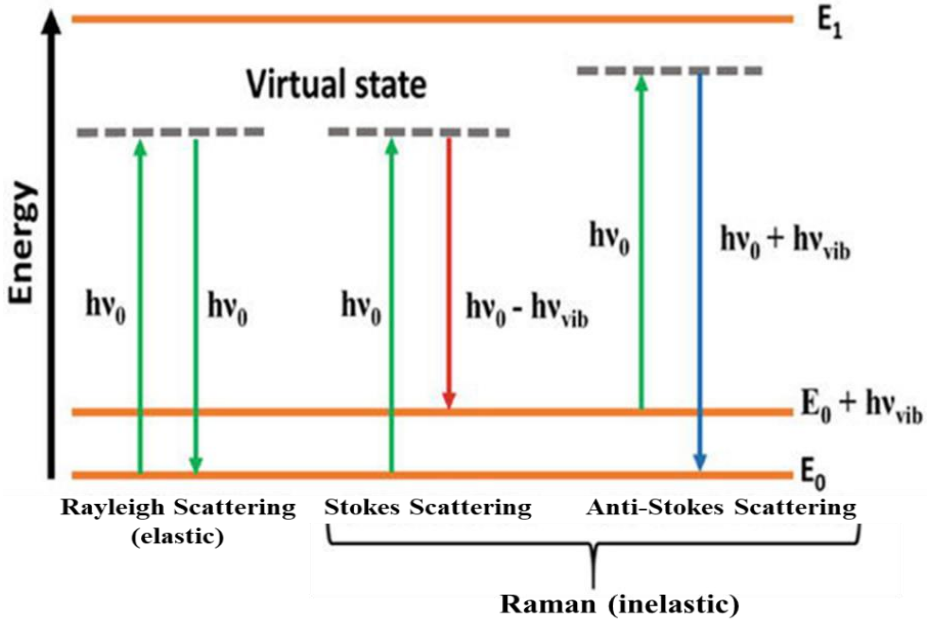

Figure 1. Energy level diagram for Rayleigh and Raman scattering

energy $\Delta E = h\nu_{vib}$, where $h\nu_{vib}$ is the difference in the vibrational or rotational energy levels of that molecule, the photon will be scattered with the energy ($h\nu_0 - h\nu_{vib}$), and the scattering is known as Stokes scattering. If the molecule loses energy $\Delta E = h\nu_{vib}$, the scattered photon will have energy ($h\nu_0 + h\nu_{vib}$), which is called anti-Stokes scattering. Fig. 1 represents the energy level diagram for Rayleigh and Raman scattering.

1.2.2 Enhancement of Raman Scattering

There are various techniques for the enhancement of Raman signals, including resonance Raman spectroscopy [25], tip enhancement Raman spectroscopy [17, 26], surface enhancement Raman spectroscopy [27], and fiber enhancement Raman spectroscopy [19, 20].

- **Resonance Raman Spectroscopy**

In the classical theory of Raman spectroscopy, the molecule's cross-section of Raman scattering depends on the frequency of the light used and the intensity of the Raman scattering signal will vary according to the fourth power of the frequency. When the frequency of the incident radiation matches the electronic transition of the irradiated molecule, the Raman scattering cross-section can be significantly enhanced by 3-4 orders in magnitude. This is the resonance Raman effect. It is possible to selectively observe a specific molecule in a complex medium by using resonance Raman spectroscopy because of the energy matching of the exciting radiation with the absorption transition [28]. A primary limitation of resonance Raman spectroscopy is that it necessitates a tunable laser, or a laser source that aligns with the vibrational frequencies of the target molecules, to achieve signal enhancement [29].

- **Tip Enhancement Raman Spectroscopy**

The scattered light from the sample is composed of propagating and non-propagating radiation. The non-propagating evanescent waves remain near the plasmon source and do not contribute to the image formation in the far field. Still, they propagate laterally on the sample among the active plasmon sites. That is why the spatial resolution of this technique is limited by the size of the focal spot of the light. This technique used a plasmonic nanotip ~ 50 nm above the sample's region of interest. The excitation light is focused onto the tip-surface cavity to induce the

localized surface plasmon resonance (LSPR) within the tip's apex and the sample surface. LSPR is a phenomenon that occurs when conduction electrons on the surface of metallic nanoparticles

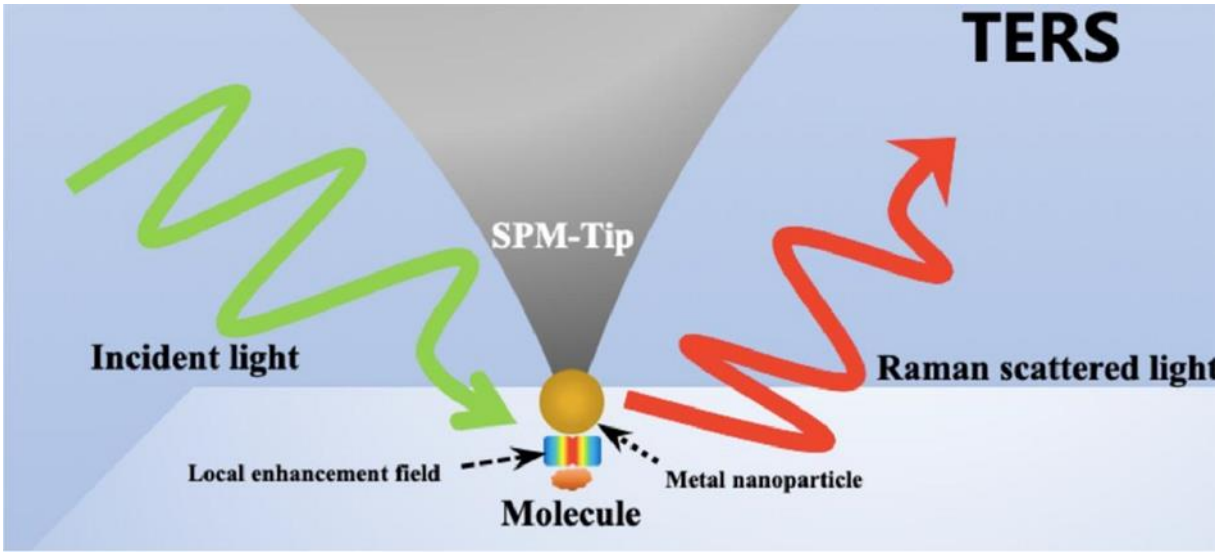

Figure 2. Schematic diagram of TERS [34].

oscillate in resonance with incident light. This oscillation creates intense electromagnetic fields near the surface of the nanoparticles, significantly enhancing the local electromagnetic environment [30]. The schematic diagram of the TERS phenomenon is shown in Fig. 2. A Raman image can be obtained by scanning the sample under the nanotip. The enhancement factor is typically in the order of 10^3 to 10^6, which is relatively smaller than the SERS due to the relative size of the probed volume. The mechanism of enhancement in this technique is like the SERS: EM and chemical enhancement. The shortcoming of the TERS is that the tip may be easily erased [17].

- **Fiber Enhancement Raman Spectroscopy**

This technique used a hollow-core photonic crystal fiber to hold a sample inside it, increasing the excitation laser's interaction volume and the sample used. The excitation laser and the scattered Raman signal are guided in the hollow core photonic crystal fiber. Since the analyte molecule is

kept in a small hollow vessel, changing the sample, and cleaning it is difficult. People have used this technique to enhance the signal from gases and antibiotics [31].

- **Surface-Enhanced Raman Spectroscopy (SERS)**

SERS is one of the amplification techniques of Raman signal from the molecules by using metallic nanoparticles or nanostructure substrates. SERS occurs mainly due to the electromagnetic interaction of light with metallic nanoparticles, which produces a strong electromagnetic field localized around nanoparticles through plasmon resonances. The first SERS enhancement was observed by Fleischmann et al. in 1974 [32]. The researcher tried to implement Raman spectroscopy to observe molecules on surfaces at a monolayer coverage during that time. They interpreted the enhanced signal due to the increased surface area of pyridine molecules due to the roughened silver electrodes. However, there were many questions unanswered. Later, other researchers claimed that the observed enhancement could not be due to the increased surface area, but other things might exist [33, 34]. Many mechanisms were proposed in the early days, but nowadays, only two are broadly accepted: electromagnetic theory (EM) and chemical enhancement (CE) theory. The CE theory explains that the enhancement is due to the chemical interaction between the probe molecules and the metal nanoparticles [35].

In contrast, the EM theory explains the enhancement due to localized surface plasmons (LSPs) [36]. The key to SERS is using plasmonic nanostructures, typically made of noble metals like gold and silver. These materials support surface plasmon resonances—collective oscillations of the electron cloud on the surface of the metal when it interacts with electromagnetic light. This interaction leads to the generation of enhanced local electromagnetic fields.

Hot spots are specific sites on the SERS-active substrate where the electromagnetic field is significantly intensified. These occur mainly at the gaps or junctions between closely spaced

nanoparticles. When the gap is in the order of a few nanometers, the electromagnetic field can be enhanced up to $10^{11} – 10^{14}$ times, dramatically increasing the intensity of the Raman signal.

The SERS has advantages and disadvantages. The benefits are label-free detection, higher enhancement factors, and detection at the single-molecule level. The main disadvantage is poor reproducibility due to the difficulty in controlling the size of nanoparticles and the position of the analyte molecules inside the hot spots [37].

- **Single Molecule SERS**

Kneip et. all first observed the Raman spectrum of crystal violet at single molecule levels using the silver colloidal solution [22]. Also, S. Nie and S. R. Emory detected the single rhodamine 6G molecule using silver nanoparticles and found an enhancement factor of $10^{14} – 10^{15}$ [38]. The intensity of surface-enhanced Raman scattering (SERS) is influenced not solely by the characteristics of the SERS substrates but also by the nature of the analytes and the optical configurations. Regarding analytes, different molecules exhibit varying Raman scattering cross-sections. Under identical excitation conditions, molecules with larger differential cross-sections produce more pronounced Raman scattering signals [39].

Detecting the Raman spectrum from just one molecule, known as single-molecule SERS (SM-SERS), is the goal for SERS sensitivity. This method is extremely sensitive and efficient for chemical analysis, allowing to see tiny details of individual molecules without averaging. SM-SERS is widely used in many areas, like chemical analysis, nanoelectronics, and biochemical sensing [40- 43]. For example, it helps monitor catalytic reactions, showing how they work at the level of single molecules and helping to make better catalysts. However, when the analyte concentration is very low, like at a single-molecule level, it brings up new and difficult issues [44]. One of the challenges in SERS is signal intensity fluctuation, which can arise from several factors,

leading to measurement reproducibility issues. These fluctuations are believed to be due to the "hot spots" phenomenon, variations in the analyte molecules' local environment, and the SERS substrates' dynamic nature. The actual reason behind the fluctuation is still unknown.

Hot spots are nanoscale regions where electromagnetic fields are significantly intensified due to the plasmonic coupling between adjacent metallic nanoparticles or nanostructures. The intensity of the SERS signal can vary dramatically depending on whether the analyte molecules are located inside these hot spots. It has been demonstrated that the spatial arrangement and the gap distance between nanoparticles significantly affect the intensity and reproducibility of SERS signals [45].

The chemical environment around the analyte molecules can also lead to fluctuations in SERS signals. The presence of other molecules, pH changes, and solvent effects can alter the analyte's orientation, adsorption kinetics, and surface coverage on the SERS substrate, thus affecting the signal intensity [46]. Another source of signal fluctuation is the physical and chemical stability of SERS substrates over time. Many SERS substrates are prone to oxidation, sulfidation, or other surface transformations that can affect their plasmonic properties and, consequently, the SERS enhancement factor [47].

To mitigate these fluctuations, many researchers have developed various strategies, including using stable and well-defined nanostructured substrates, controlled preparation methods, and advanced statistical analysis techniques to account for variability in the data. The development of reproducible and stable SERS platforms remains an active area of research, aiming to harness the full potential of this powerful spectroscopic technique.

- **Vacuum Enhancement Raman Spectroscopy (VERS)**

The vacuum enhancement Raman spectroscopy works on the principle of vacuum evaporation. When the pressure is below atmospheric pressure, the solvents evaporate, which

increases the concentration of analyte molecules. Focusing the laser on the bottom of the sample holder gives an enhanced signal.

The enhancement of the VERS depends on the concentration of the analyte molecule before and after the vacuum evaporation. If c_v is the analyte concentration before vacuum evaporation and c_o is the concentration after vacuum, the enhancement factor will be c_v/c_o. For a given initial volume V and the depth of laser focal volume of $2z_R$, the enhancement factor is given by

$$EF = \frac{V}{\pi r^2 * 2z_R} \tag{10}$$

Where $z_R = \frac{\pi W_0^2 n}{\lambda}$ is the Raleigh distance, w_0 is the beam waist, λ is the wavelength, and n is the index of refraction.

1.2.3 SERS Substrates and Enhancement Mechanisms

A suitable SERS substrate must be able to provide substantial enhancement of the Raman signal. The most common substrates used are gold and silver nanoparticles. The gold and silver nanoparticle works the best due to their optical properties, specifically in the visible and near-infrared range. The nanostructure is the conformation that increases surface area to ensure the contact between the metal and substrate. Roughened nanoparticles are beneficial due to the even higher surface area [48].

- **Gold Nanoparticles**

Gold nanoparticles (Au-NPs) are easy to make and only require a few ingredients. There are many ways of synthesizing gold NPs, and the widely used method is the Turkevich method [49]. This method is rapid and reliable, by using Chloroauric acid as the precursor and trisodium citrate dihydride as the reducing and stabilizing agent. The size of the gold NPs can be controlled by the amount of sodium citrate in the solution. The nanoparticle size was estimated using an electron

microscope. Also, UV-VIS spectroscopy is used to correlate the size of the nanoparticles [50]. Gold nanoparticles are favored over other metal nanoparticles due to several reasons including optimal surface plasmon resonance in visible to near infrared range, chemically stable, and biocompatible.

- **Electromagnetic Theory**

Most SERS enhancement comes from the EM enhancement due to the metal nanoparticles' roughness. This effect comes in two ways due to the roughness present on the nanoparticle surface. The first is the local field enhancement, and the second is the radiation enhancement.

- **Local Field Enhancement**

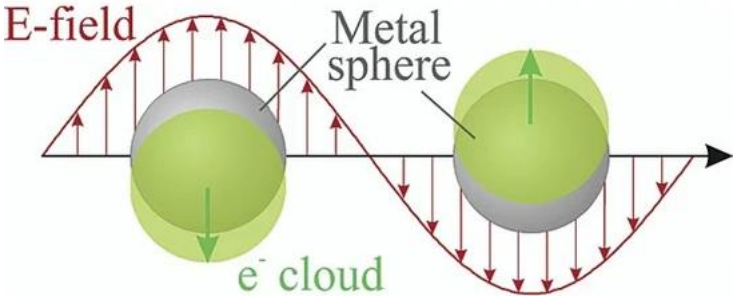

Figure 3. Origin of electromagnetic field enhancement mechanism in SERS [11].

When electromagnetic radiation strikes metal surfaces, the electromagnetic field in the vicinity of metallic surfaces gets strongly modified due to the coherent oscillation of free electrons generating surface plasmons. There are two types of surfaces plasmons – propagating plasmons and surface plasmons. Propagating plasmons propagate in x- and y-direction along with the metal-dielectric interface [51]. In contrast, the surface plasmon is localized on the surface of the nanoparticle with a frequency known as the localized surface plasmon resonance (LSPR) [52, 53]. The local field E_L on the metallic object at the molecular position could differ from the incident

field E in terms of magnitude and direction, such that the magnitude of $|E_L|$ is greater than |E|. The Raman dipole induced by the local field is

$$\mu_R = \alpha_R E_L(\nu_L) \tag{11}$$

The Raman dipole is then enhanced by a factor of $|E_L(\nu_L)|/|E|$. The local electric field enhancement factor $M_L(\nu_L)$ over the absence of a metallic object is proportional to $|\mu_R|^2$, which is

$$M_L(\nu_L) = \frac{|E_L(\nu_L)|^2}{|E|^2} \tag{12}$$

- **Radiation Enhancement**

Because of the presence of the metallic object in SERS, the metallic environment significantly alters dipole radiation in the case of the incident electric field. The total power radiated by the dipole, P_{rad}, can be either enhanced or quenched (relative to free space P_0) depending on the relative dielectric function of the object, its geometry, and the dipole position, orientation, and the emission frequency. When the dipole is parallel to the surface of the non-absorbing dielectric sphere, like glass, $\varepsilon(r) > 1$, small quenching in P_{rad} is observed. For objects having negative real dielectric function with $\mathrm{Re}(\varepsilon(r)) < 0$, the radiated power is strongly enhanced. The larger enhancement is due to coupling to the LSPR of metallic objects [54]. The radiation enhancement factor is therefore given by

$$M_{rad} = \frac{P_{rad}}{P_0}. \tag{13}$$

The overall enhancement factor of local field enhancement and radiation enhancement is $M_{tot}=M_L(\nu_L)*M_{rad}$.

- **Chemical Theory**

This is the second enhancement mechanism considered independent of the EM enhancement. This is considered a short-range chemical effect in which the polarizability of the adsorbed molecules changes. The chemical enhancement (CE) mechanism is viewed as resonance polarizability of the adsorbate due to the formation of the metal adsorbate complex. The CE can be grouped into three contributions: (i) a resonance Raman (RR) effect due to the incident light matching an electronic transition in the molecule, (ii) a charge transfer (CT) effect where the incident light is in resonance with a metal molecule or molecule-metal transition, and (iii) a non-resonant chemical effect due to ground state orbital overlap between the molecule and the metal [55]. RR is a molecular resonance mechanism that arises when the incident light is resonant with a molecule and leads to the resonance Raman scattering. This does not require any metal particles. The CT effect appears when the metal particles and the molecules are close enough to allow for a sufficient wavefunction overlap. The adsorption of the analyte on the metal surface results in the formation of a metal-adsorbate complex, and due to the covalent bonding between adsorbate and metal surface, a new electronic state arises and serves as a resonant intermediate state. That means the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO) of the adsorbate are symmetrically disposed of energy with the Fermi level of the metal. In this way, the charge transfer occurs from the Fermi level to LUMO and retro donation of an electron from HOMO to the Fermi level of the metal. This transfer of electrons creates negative ion formation. Enhancement occurs when the energy of the negative ion is resonant with the incident light. The charge transfer effect is short-range and strongly dependent on the geometry, bonding, and molecular energy level [56]. The non-resonant chemical effect is the least studied effect and most difficult to quantify because of the low contribution to the overall enhancement.

The formation of molecular metal complexes is the leading cause of this effect due to chemical bonding [57].

1.2.4 Raman Imaging

Raman imaging, a vibrational spectroscopic technique rooted in the Raman effect, merges Raman spectroscopy with optical microscopy to create high-resolution images of analyte molecules [58]. Utilizing a CCD camera, which comprises light-sensitive detectors, Raman imaging converts incident light into electrical charges stored in pixels corresponding to different Raman spectrum wavelengths. These charges reflect light intensity, constructing a detailed Raman spectrum image across the sample.

There are many ways to capture the Raman images. One method is by scanning a laser across the sample and capturing Raman spectra at each point. The other method is irradiating the sample volume and recording the Raman images of the irradiation volume by a CCD camera [59]. Raman imaging forms an image mapping chemical variation, offering insights into microscopic and nanoscopic physical structure and chemical properties. Its strength lies in providing detailed molecular and structural insights applicable in diverse fields like material design, disease diagnosis, and artifact analysis. CCD cameras record the spatial distribution of Raman signals, translating them into digital images where each pixel represents a sample position while holding Raman scattering intensity data. Analyzing Raman intensity at each pixel generates a molecular information map, revealing chemical species distribution and concentration, aiding understanding of sample composition, structure, and spatial diversity [60]. In summary, Raman imaging with CCD cameras provides a potent tool for spatially resolved molecular analysis, offering insights into sample compositions, structure, and distribution in field like material science and biology.

1.2.5 Cell, Nucleus, Amino acids, and Proteins

- **Cell**

A cell is a basic membrane-bound unit that contains the fundamental molecules of life. Some cells become specialized when they mature and perform different functions. All the cells cooperate with the specialized cell and become the building block of multicellular organisms such as humans. Although the cells are much larger than the atoms, they are still tiny. As an individual cell, it processes its nutrients, synthesizing many types of molecules, providing energy, and replicating itself to produce new generations. In a multicellular organism, different cells communicate with each other. Similar kind of cells form tissue and then forms organs [61].

A cell is enclosed by a cell membrane, forming a selective barrier that allows nutrients to enter the cells and waste products to leave. It also protects the cell from outside. The inside cell

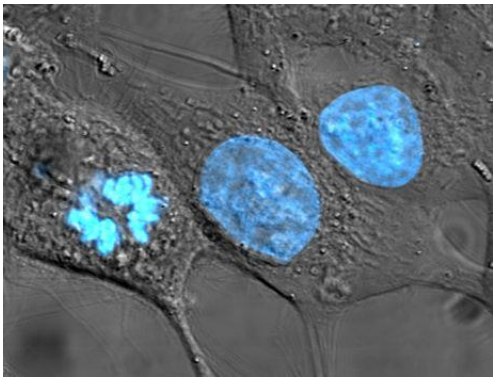

Figure 4.Human cancer cells, HeLa cells, with DNA stained blue.

contains many organelles, each surrounded by a separate membrane. One major organelle is a nucleus, which contains genetic information. Mitochondria is another organelle responsible for the energy transactions necessary for cell survival. Between the organelle there is the cytoplasm, which contains cytosol. The cytosol contains an organized framework of fibrous molecules, which gives a cell its shape. The RAS protein is attached to the cell membrane [61].

- **Nucleus**

The nucleus is the core component of a cell, and it mainly contains protein and DNA. The DNA molecules are organized in a long chain called chromosomes. The chromosomes contain the functional unit called a gene. The protein inside the nucleus belongs to the specific type and provides essential support for the DNA. An outer membrane to the nucleus is like a cell membrane called a nuclear membrane. Inside the nucleus, there is a nucleolus, the RNA-rich region, which helps with gene transcription [61].

- **Amino acids**

Amino acids are the monomers, and each amino acid has the same basic structure: a carbon atom linked with an amino group (NH2), a carboxyl group (COOH), one hydrogen atom, and an

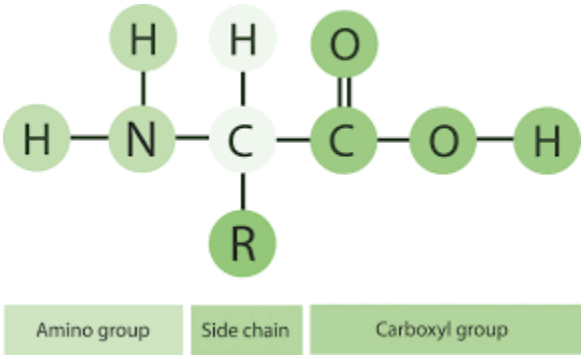

Figure 5. Amino acids.

R group, also called a side chain. There are 20 amino acids, and proteins are formed using these amino acids. The main essential amino acids are alanine, valine, tyrosine, tryptophan, and histidine [61].

- **Proteins**

Proteins are the organic molecules found in living systems, and they have a diverse range of functions. Proteins are long-chain molecules of amino acids joined together by peptide bonds. All

amino acids are compounds with a similar basic structure. It has an amino, carboxylic, and sidechain or R group specific to each protein. The protein variation is due to the order of the amino acids, chain length, and chain folding in different shapes. There are different proteins in the human body, and we will study the RAS proteins in this work [61]. The protein has different primary, secondary, tertiary, and quaternary structures. The primary structure is simply the sequence of the amino acids that determines the protein's shape and function. The secondary structure describes the localized shape of the proteins. The secondary structure contains an alpha helix, beta-sheet, random coil, and beta-turn. Hydrogen bonds stabilize the alpha helix structure. The N-H group of one amino acid interacts with the carbonyl group of another amino acid. The beta-sheet is also

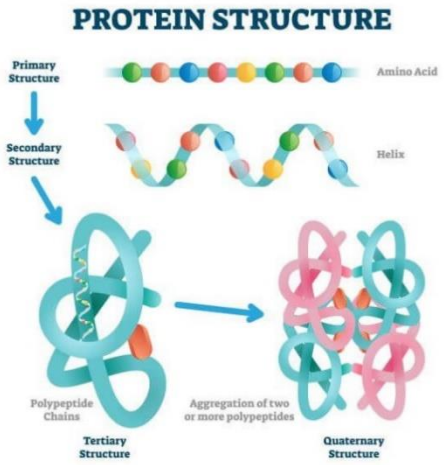

Figure 6. Protein structures.

stabilized by a hydrogen bond formed by the N-H group of one amino acid interacting with the carbonyl group of another amino acid. The tertiary structure is the three-dimensional complete folded pattern of the protein, which is one individual subunit. The quaternary structure is the multiple subunits [62].

Raman spectrum of protein can provide detailed insights into its molecular structures and dynamics. Different molecules in protein contribute to different bands in the Raman spectrum. Table 1 gives the assignments of different bands of protein.

Table 1. Band assignments for Raman spectrum of protein.

Raman shift (cm^{-1})	Tentative assignments	references
418	Trp	[63]
449	ν(C-S)	[64]
465	ν(C-S)	[64]
498	ν(S-S)	[68]
512	ν(S-S)	[68]
535	δ(N-H), ν(S-S), δ(skeletal)	[65], [66]
550	ν(S-S)	[66]
575	ν(S-S)	[69]
601	$\delta(COO^-)$	[67]
642	ν(C-S)	[68]
655, 662	ν(C-S), Tyr	[68], [69]
682	δ(-C-H)	[69]
694	C-C, C-O bend	[68]
706	ν(C-S), δCOO^-)	[68]
717	ν(C-S), Trp	[68]
740	Trp, ν(C-S)	[68]
772,779	Trp, δ(-C-H)	[64]
797	δ(C-H), δ(-N-H)	[68], [64]
838	Tyr	[68], [64]
850	Tyr:R breathing	[68], [69]
868	Tyr	[70]
895	Trp	[68]
940	δ(-C-C-N) symm, α-helical skeletal	[68]
958	ν(N-C_α-C) skeletal	[68]
1004	Phe	[68], [64]
1049,1059	ν(C-N), Phe	[68], [64]
1094	ν(C-C), ν(C-N), H(r) bend	[68]
1132	δ(-C-N), Pro	[68]
1166	Tyr	[68]
1175	Tyr, ν(-C-N)	[68], [64]
1210,1218	Tyr, Phe, ν(C-C)	[64], [69]

1252	Amide III	[68], [64]
1295	Amide III, C-H, C-C	[66], [69]
1324	δ(C-H), $\delta(CH_2/CH_3)$	[68], [64]
1344	$\gamma(CH_2, CH_3)$, Trp(C_α-H)	[66], [64]
1354	$\gamma(CH_2)$	[64]
1369	$\delta(CH_3)$ symm	[70]
1410	$\delta(CH_3)$	[68], [64]
1425	$\delta(CH_3)$Cys	[68]
1434	$\delta(CH_2)$	[67]
1444	δ(C-H) of CH_2	[66]
1458	δ(C-H), $\delta(CH_2/CH_3)$	[68], [66]
1504	Amide II, ν(R,r), Trp, ν(C-H)	[64]
1526	$\delta(NH_3^+)$Lys, amide II	[64]
1555	Trp: ν(R), ν(r), amide II	[65]
1565	Amide II	[64]
1577	Trp, NH_2 sciss, ν(R,r)	[65]
1586	δ(R), ν(R)	[64]
1595	Phe, Tyr: ν(R), $\nu(COO^-)$asymm, OH mode	[65]
1629	Amino acids	[66]
1654	Amide I	[66]

- **RAS proteins**

The RAS proteins are members of the small GTP binding and hydrolyzing proteins (GTPases). They are the signal transduction regulator for cell proliferation, survival, and differentiation [2]. These proteins are the GDP/GTP binding guanine triphosphatase encoded by the Ras gene [61]. These proteins cycle between the GTP (active state) and GDP (inactive state) bound states and act as a binary molecular switch that functions in a signaling event.

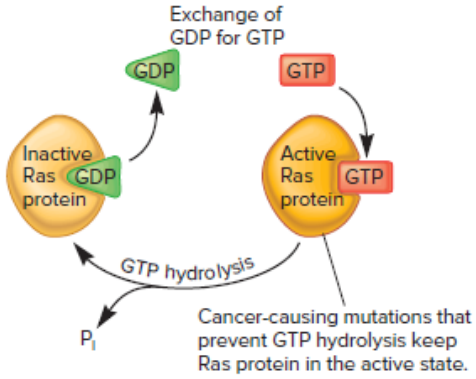

Figure 7. The functional cycle of Ras protein. GTP binding to Ras activates its function and promotes cell division. Conversely, the hydrolysis of GTP to GDP and Pi converts the active form of Ras to an inactive form.

Ras proteins usually are tightly regulated by guanine nucleotide exchange factors (GEFs), promoting GDP dissociation and GTP binding and GTPase-activating proteins (GAPs) that stimulate the intrinsic GTPase activity of Ras to switch off signaling [72]. This is responsible for the functional interactions of these proteins with negative (GAP) and positive (GEF) cellular regulators [73]. The main known function of the RAS protein is mediating extracellular proliferative signals controlled by receptor tyrosine kinases (RTKs) in response to growth factors to a variety of effector molecules stimulating a cascade of parallel phosphorylation reaction pathways that finally activate nuclear transcription factor [74]. During the regular operation of RAS, it performs its function in a GTP binding form. The RAS has low hydrophobicity, and hence

it cannot bind with the cell membrane itself, so it must undergo some modification with enzymes. This increases the hydrophobicity and binds to the inner cell membrane. Once the RAS performs its function in GTP binding sites, it hydrolyzes into the GDP binding form [71]. Oncogenic

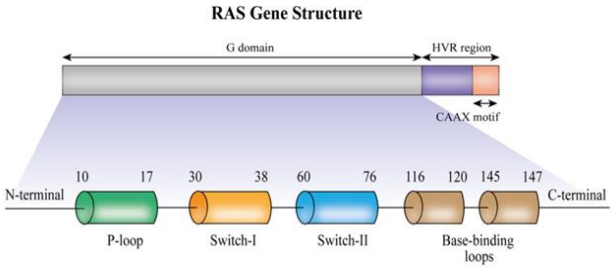

Figure 8. Ras gene structure. The G-domain corresponds to the initial 166-168 residues, which contain the phosphate-binding loop (P-loop, residues 10-17), switch I (residues 30-38), switch II (residues 60-76), and the base binding loops (residues 116-120 and 145-147). The carboxy terminal contains a hypervariable region (HVR), including CAAX, where C is a cysteine amino acid, A is an aliphatic amino acid, and X is an amino acid.

mutations at positions 12, 13, or 61 of the HRAS, NRAS, and KRAS genes are the most common genetic mutations in the tumor cells. This kind of mutation results in the significant impairment of the GTPases activity of the RAS protein and locks them into the active state even in the absence of stimuli, which leads to the formation of uncontrolled cell growth and finally leads to the tumor [73]. The 3D crystal structure of kras4b (1-169) [75] and HRAS (1-189) [76] are shown in Fig.

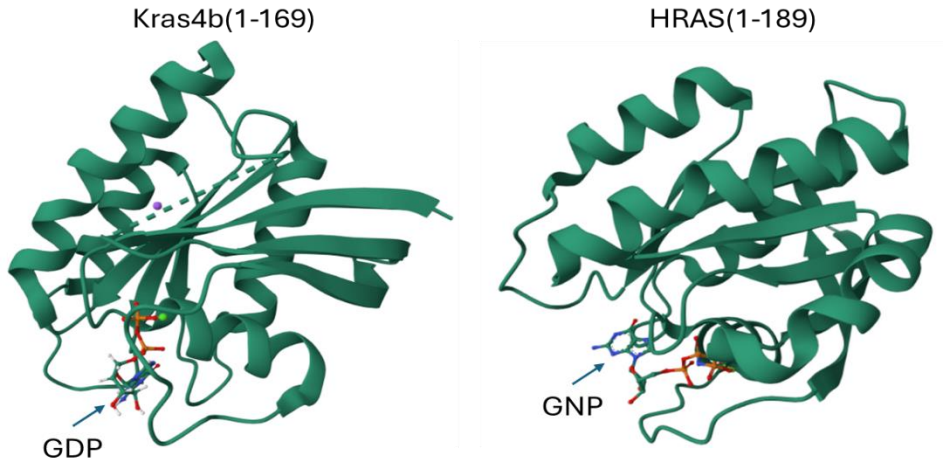

Figure 9. Crystal structure of kras4b (1-169) and HRAS (1-189).

9. The 3D structure shows the secondary structures of protein. It consists of a coiled-like structure called an alpha helix. The sheet-like structure is called a beta sheet. The coil connecting the alpha helix and beta sheet is the random coil. For kras4b (1-169), GDP molecules are also attached, and for HRAS (1-189), GNP molecule is attached.

- **ARS-1620 drug compound**

ARS-1620 is the KRAS G12C inhibitor synthesized from the quinazoline series. For many years it was believed that the KRAS mutation was undruggable because of its small size, and strong binding to GTP. But, after discovering the switch II pocket of the KRAS, it was discovered that the G12C version of KRAS can be druggable with small inhibitor molecules such as ARS-1620. Scientists are trying to use this compound to suppress the growth of mutant KRAS. It was discovered that the ARS-1620 drug binds covalently with the cysteine-12 of the KRAS G12C proteins in vitro [77].

1.3 Scientific Questions, Objectives of Present Work, and Significance

- **Scientific Questions to Address**

There are three main scientific questions that I will address in this research work. They are as follows.

a) Does normal Raman spectroscopy help to differentiate RAS isoforms?

RAS proteins act as binary molecular switches that regulate signaling pathways within cells. Mutations in these proteins can lead to uncontrolled cell growth and tumor formation. It has been observed that various cancers are associated with mutations in specific types of RAS isoforms. This research will first investigate whether normal Raman spectroscopy can differentiate between various RAS isoforms. Since RAS proteins operate between GDP and GTP states, the study will

also assess whether Raman spectroscopy can identify these distinct states of GDP and GTP binding. Techniques such as Principal Component Analysis (PCA) and Discriminant Analysis of Principal Components (DAPC) will be used to differentiate the RAS spectra obtained from Raman spectroscopy. Additionally, this project aims to predict the secondary structure of RAS isoforms through numerical fitting of the amide I band in the Raman spectrum. The hydrophobicity and hydrogen bonding conditions of different isoforms will also be examined. Lastly, this research will explore whether Raman spectroscopy can detect the interaction of the small inhibitor drug molecule ARS-1620 with the KRAS G12C isoform.

b) Can we develop new Raman enhancement techniques for low concentration liquid sample?

Various Raman enhancement techniques have been developed to detect molecules in solution, but these methods often enhance the signal from both the analyte and the solvent molecules. This can be particularly disadvantageous when the low concentration analyte molecules are dissolved in solvents with a high Raman background. The objective here is to develop a technique that specifically enhances the signal from analyte molecules only at low concentration, reducing interference from the solvent.

c) Can we use SERS and Raman imaging for the RAS analysis at nanomole concentration level?

SERS is highly enhanced technique with large enhancement factor. This uses the metal nano particle such as silver or gold nanoparticle to enhance the signal. This nanoparticle can detect the analyte signal at single molecule levels. The Raman imaging technique, particularly when enhanced by SERS, offers a highly sensitive method for visualizing molecular compositions at high resolution. It allows for monitoring Raman spectral changes over time or under varying conditions, such as environmental shifts or molecular interactions. By combining with SERS,

Raman imaging gives insights into the dynamic chemical and physical behaviors of the RAS protein, making Raman imaging a crucial tool in fields like biology, materials science, and chemistry.

- **Objectives of Present Work**

The objectives of this dissertation are as follows.

1) To apply the normal Raman spectroscopy technique to characterize the RAS protein isoforms.
2) To develop a new technique to enhance the Raman signal of low concentration liquid samples.
3) To observe the interaction of ARS-1620 molecules with KRAS G12C RAS isoforms.
4) To develop SERS and Raman imaging techniques particularly for RAS protein study.

- **Significance**

Raman spectroscopy is a powerful tool for the non-invasive detection of molecules. However, it suffers from low scattering efficiency and due to this it is challenging to detect the molecules for low concentration samples. Enhancement techniques for low concentration samples is very important to detect such molecules.

Significant research efforts are ongoing worldwide in cancer studies, focusing particularly on the dysfunction of RAS proteins, which are often implicated in various cancers. Understanding and identifying different RAS isoforms is crucial for the scientific community to gain deeper insights into cancer mechanisms. Additionally, as scientists pursue targeted therapies for cancer

treatment, it becomes increasingly important to identify specific properties of RAS isoforms—such as hydrogen bonding conditions, hydrophobicity, and secondary structure.

The interaction between the small molecule ARS-1620 and KRAS G12C proteins is particularly noteworthy. Studying this interaction helps understand the effects of ARS-1620 on the proteins, which could inform future uses of the drug molecule for controlling mutant RAS and, consequently, cancer.

Studying the secondary structure is essential for understanding the native and mutant states of proteins or their stable and unstable states. The more alpha-helix there is in the secondary structure, the more stable the protein. Moreover, analyzing the peak position of the amide I band of the protein indicates the ordered or disordered structure of the proteins. If the center of the amide I peak lies between 1650-1658 cm^{-1}, it indicates a high alpha-helical content. If it lies between 1660-1665 cm^{-1}, it suggests a high proportion of random coil or disordered structures. Additionally, studying the hydrogen bonding condition and hydrophobicity of different RAS isoforms provides insights into the dynamics of the proteins, which is crucial for understanding protein-drug interactions.

The concentration of Ras inside cell is in nanomole range which requires highly enhanced technique such as SERS to detect. SERS can detect signal at single molecule levels. Also, SERS based Raman imaging is a highly significant technique in both scientific research and industrial applications due to its unique capabilities in chemical characterization and spatial resolution. It provides detailed information about the molecular composition of sample without the need for labels. This is particularly valuable for studying biological samples.

1.4 Literature Review

Brochnike et al. studied the structural changes by fluorescence spectroscopy when Ras is binding with GDP, GTP, or non-hydrolyzable analogs GppNHp, $GppCh_2p$, and GTP-γS nucleotide to distinguish between the active and inactive form [78]. By looking at the emission spectra, they distinguished the active and inactive state of Ras, as shown in Fig. 10, where the emission spectra for the inactive state (GDP) is higher than the active state because of the different exposure of the tryptophan residue to the solvent.

Maiti et al. applied the peak fitting technique to the Raman amide I band of α-synuclein proteins to characterize the secondary structures [79]. They were able to predict the secondary structures of different proteins, and they found that these proteins are natively unfolded as the center of the amide I band has shifted to the 1670 cm^{-1} region as shown in Fig. 11.

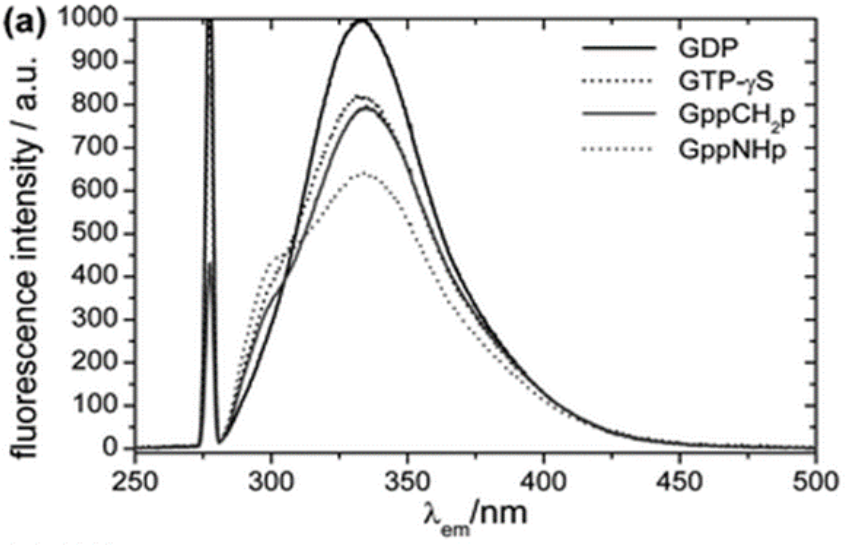

Figure 10. Fluorescence properties of RasY32W bound to GDP, GTP-γS, GppCh2p, and GppNHp.

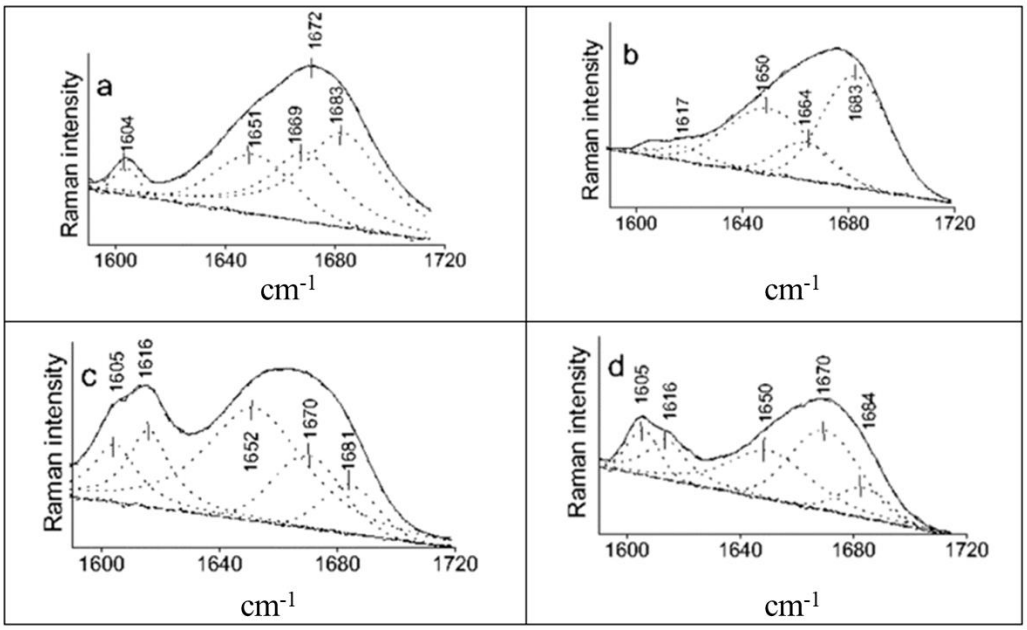

Figure 11. Curve fitting of amide I band of a) NAC, b) phosvitinn, c) α-casein, d) β-casein.

Wang et al. studied the vibrational spectra of phosphate modes for GDP and GTP binding of HRAS using Raman spectroscopy and ab initio calculations [80]. Their experiment showed that the phosphate stretch frequencies are different, suggesting the different conformations when Ras is in the GDP and GTP state. Liu et al. distinguished between the wild-type and mutant KRAS gene in colon cancer using Raman spectroscopy and analyzed with PCA and linear discriminant analysis [81]. They found that the index of correct classification was 81.2 % between wild-type and mutant KRAS.

Fig. 12 shows that the PCA model can differentiate the wild type cells from mutant cells. Signorelli et al. used the Raman spectroscopy technique to study the effect of azurin drug in P53. They focused on the tyrosine and tryptophan residues, which are the diagnostic markers of protein side-chain environment and amide I band to quantify protein secondary structure. They found evidence of structural changes within p53 protein upon interaction with azurin. They also observed increased hydrophobicity when binding with the drug azurin. This suggests that Raman

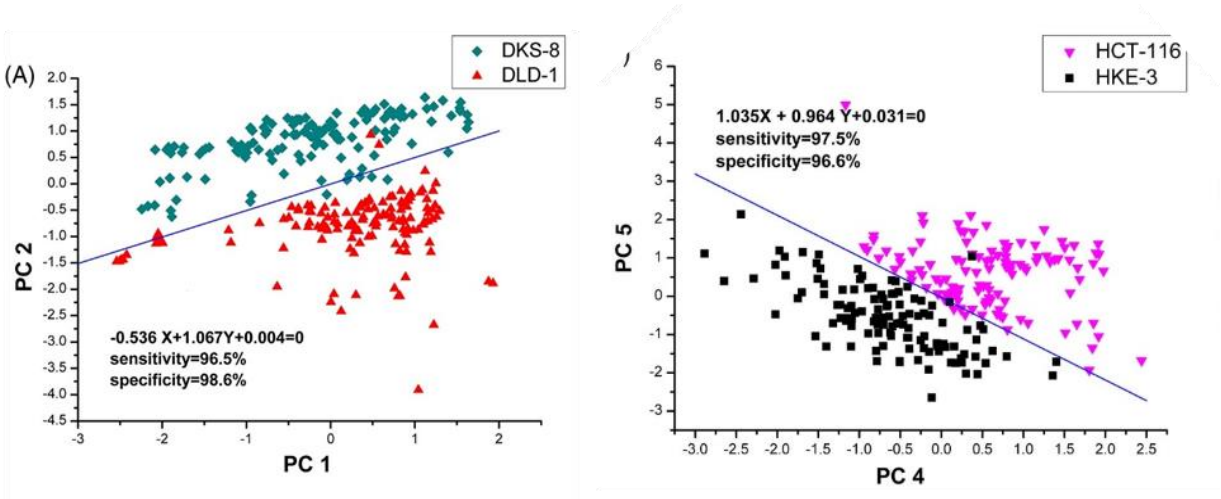

Figure 12. Scatter plots of the first principal component (PC1) vs. second principal component (PC2) for DKS-8 (KRAS wild type) vs. DLD-1 cells (KRAS MT) (left). Scatter plots of the fourth principal component (PC4) vs. fifth principal component (PC5) for HCT-116 (KRAS MT) vs. HKE-3 cells (KRAS WT) [81].

spectroscopy can study the structural and conformational changes in the protein when binding with drugs [82] SERS has then been employed for the highly sensitive detection of different proteins like flavoproteins [83], hemoglobin [84], myoglobin [85], etc. Nabiev et al. [86] have studied the SERS of aromatic amino acids and proteins on the surface of the silver hydrosols. Kneipp et al[57] demonstrated SERS in living cells, opening the possibility of in vivo and real-time applications. It was reported that an enhancement factor of 10^2-10^{14} was observed for SERS. The main advantage of SERS is a large enhancement, but it suffers many drawbacks for protein characterization. Fig. 13 shows the shape and size of various proteins. The major problem in the SERS detection of proteins is the diversity of proteins in size, shape, and surface charge so that the enhancement factor is protein- and nanoparticle-dependent. Also, the SERS has poor reproducibility due to the random protein immobilization on SERS substrates. Since the protein is a big molecule compared to other small molecules, only the part of the protein meets the nanoparticle or inside the gap region of the hot particle. Hence, the nanoparticle only enhances a part of proteins, creating different

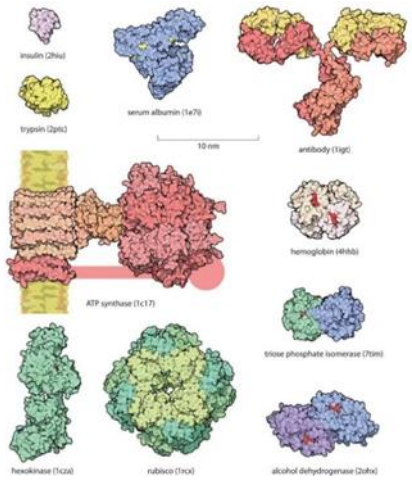

Figure 13. Different shapes and sizes of various proteins.

spectral characteristics than the bulk protein spectrum. This creates difficulty in analyzing the SERS spectrum for characterizing proteins. The nanoparticles also play an important role in the detection process. Different nanoparticles have different surface plasmon resonance, which makes

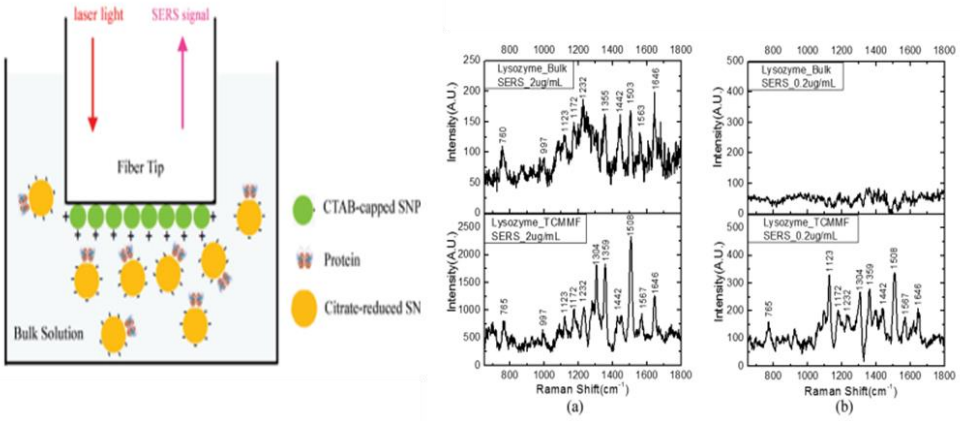

Figure 14. Schematic of the TCMMF SERS probe in aqueous protein detection(left) and SERS spectra of lysozyme detected by bulk solution and TCMMF probe at (a) 2 μg/ml; (b) 0.2 μg/ml (right) [87].

the enhancement of the Raman signal different from particle to particle. The enhancement also depends on the shape and size of nanoparticles, and it is also challenging to control such properties. Yuan et al. applied the fiber crystal and nanoparticles to detect the lysozyme and cytochrome c protein. The technique they used, and spectra are shown in Fig. 14. They achieved the lowest

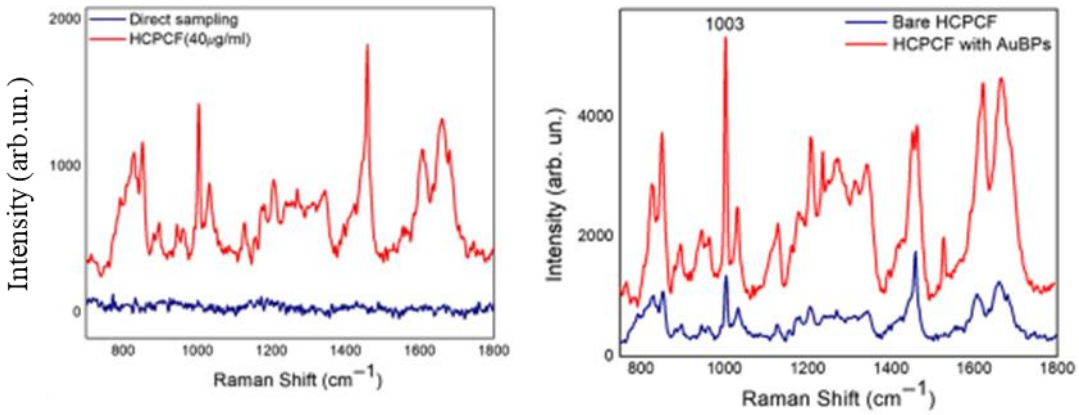

Figure 15. Raman spectra of 40 μg/ml Aβ: for direct measurement and with crystal fiber (left), fiber crystal, and fiber crystal combined with gold particles (right).

detection concentration limit of 0.2 μg/ml [87]. The hollow core photonic crystal fiber was used to detect Alzheimer's disease biomarkers amyloid β-peptide (Aβ) and tau protein, and a 20-fold enhancement factor was observed. In addition, they combined this crystal fiber with gold nanoparticles to achieve a 200-times overall enhancement.

To overcome the limitation of the diffraction limit in the detection of nanosized molecules, M. Ovesný et al. [82] introduced an imaging processing technique based on ImageJ called ThunderSTORM. They developed this technique to analyze the images obtained from Single-Molecule Localization Microscopy (SMLM). Fig. 15 is taken from M. Ovesný et al. [82]. This shows that ThunderSTORM can show more detail on a smaller scale.

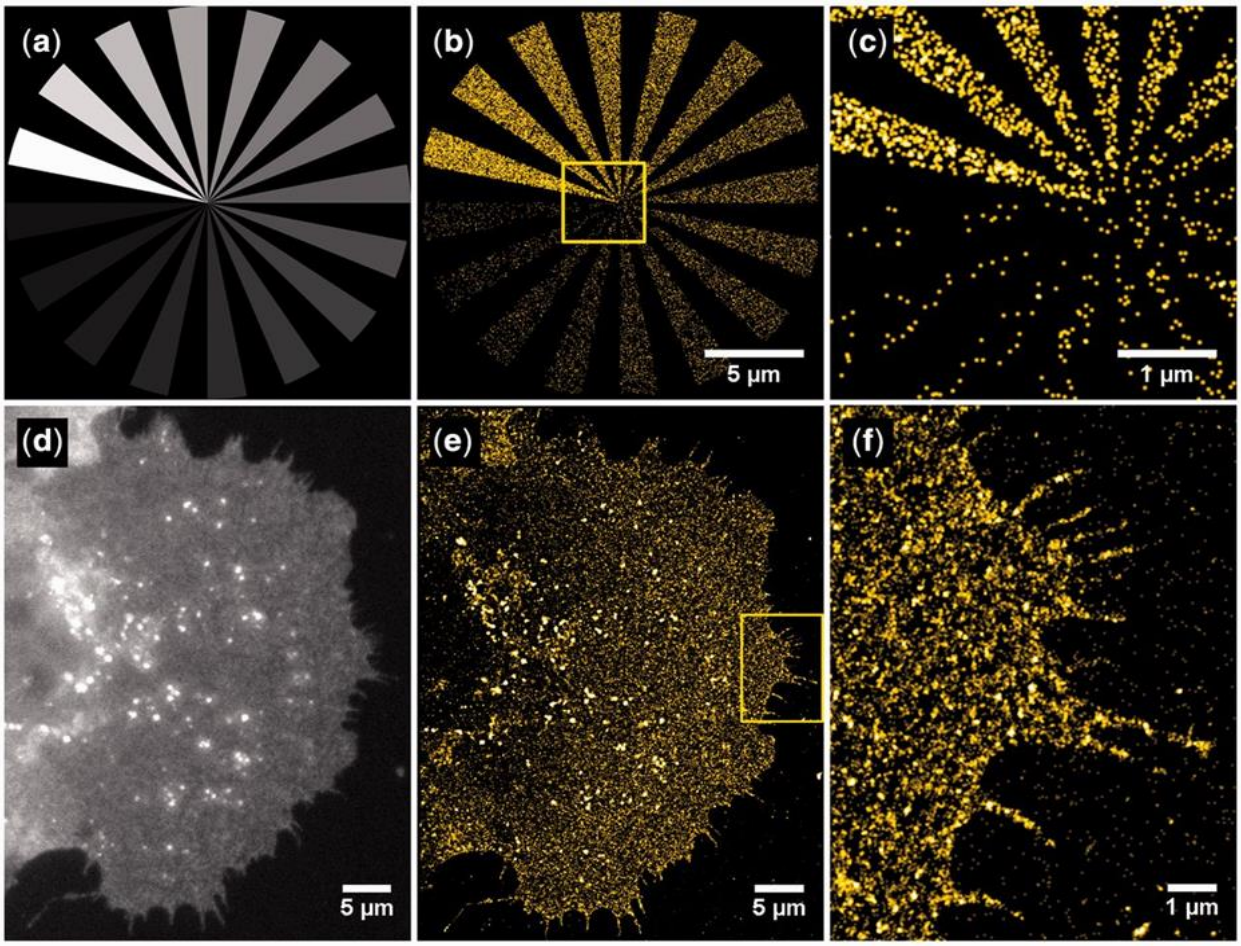

Figure 16. Simulations and SMLM reconstruction with ThunderSTORM. (a) Example of a mask used for generating simulated SMLM data. The gray-scale intensity values are interpreted as molecular densities within a user-specified range. (b) SMLM reconstruction of a simulated dataset. (c) Detail of b. (d) Widefield fluorescence image of an A431 epidermoid carcinoma cell expressing the membrane protein mCitrine-erbB3. (e) SMLM reconstruction using the default settings. (f) Detail of e [82].

1.5 Dissertation Outline

This dissertation is organized in the following. The current chapter gives the brief introduction about the RAS protein, Raman spectroscopy, Raman enhancement techniques, objective, and significance of this research work. Chapter 2 discusses the specifics of Raman instrumentation and data analysis techniques. Chapter 3 presents the normal Raman spectrum library for GDP and GppNHp Ras isoforms. Discriminant analysis of principal components

(DAPC) will be used to differentiate among these RAS isoforms. Additionally, a model based on principal components is developed to identify unknown protein samples. Deconvolution of Raman spectrum enables predictions about various characteristics such as hydrogen bonding conditions, hydrophobicity, and protein secondary structure. A Raman barcode is also generated from the deconvoluted bands. The chapter concludes with an examination of the effectiveness of Raman spectroscopy in studying the interaction of ARS-1620 with kras4b G12C. Chapter 4 explores the use of vacuum enhancement Raman spectroscopy to amplify the Raman signal. The effectiveness of this technique will be tested using rhodamine B, and it will be subsequently applied successfully to BSA, glucose, ciprofloxacin, and ARS-1620 molecules. Chapter 5 focuses on the SERS and Raman imaging study for the potential application to ras protein. Finally, the Chapter 6 provides the summary and conclusions of this research work.

Chapter 2. Raman Spectroscopy Instrumentation and Data Analysis

2.1 Raman Spectroscopy Instrumentation and Measurement Setup – 780 nm System

The schematic setup of the Raman system is shown in Fig. 17. The system's main components include a microscope, diode laser, spectrograph, and other optical components like a lens, mirror, etc. The excitation wavelength of 780 nm from a semiconductor laser (TEC-300-0780-0500, Sacher Lasertecknik Inc.) coupled into a single-mode optical fiber is used. The laser source has a frequency bandwidth of fewer than 1 MHz, and a narrow bandpass filter (LL01-780-12.5, Semrock Inc.) is used to purify the laser beam spectrally. A pair of lenses are used to expand the laser beam to fill the large numerical aperture of the microscope's objective. The laser beam is then introduced into the objective lens of the inverted microscope through the dichroic mirror (DM). The microscope used is the Nikon TE-2000S, and the objective is Nikon 100X having a 1.3 numerical aperture (NA). A dichroic beam splitter (LPD01-785RU-25, Semrock Inc.), labeled as HNF1 in Fig. 17, is used to reflect a 780 nm excitation beam and transmits the Raman signals above 810 nm. The scattered light from the sample is collected with the same objective and passes through the long wave pass filter and is focused onto the entrance of an imaging spectrograph. A long wave pass filter (LP01-780RS-25, Semrock Inc.), labeled as HNF2 in Fig.17, is used in the front of the spectrograph to further reject the elastic scattering at 780 nm. A video camera is also used for visualization. Blue LED light is used for illumination. We used the Triax 320 spectrometer (Jobin Yvon Inc) and PIXIS 100BR CCD camera (Princeton Instruments) to record Raman spectra.

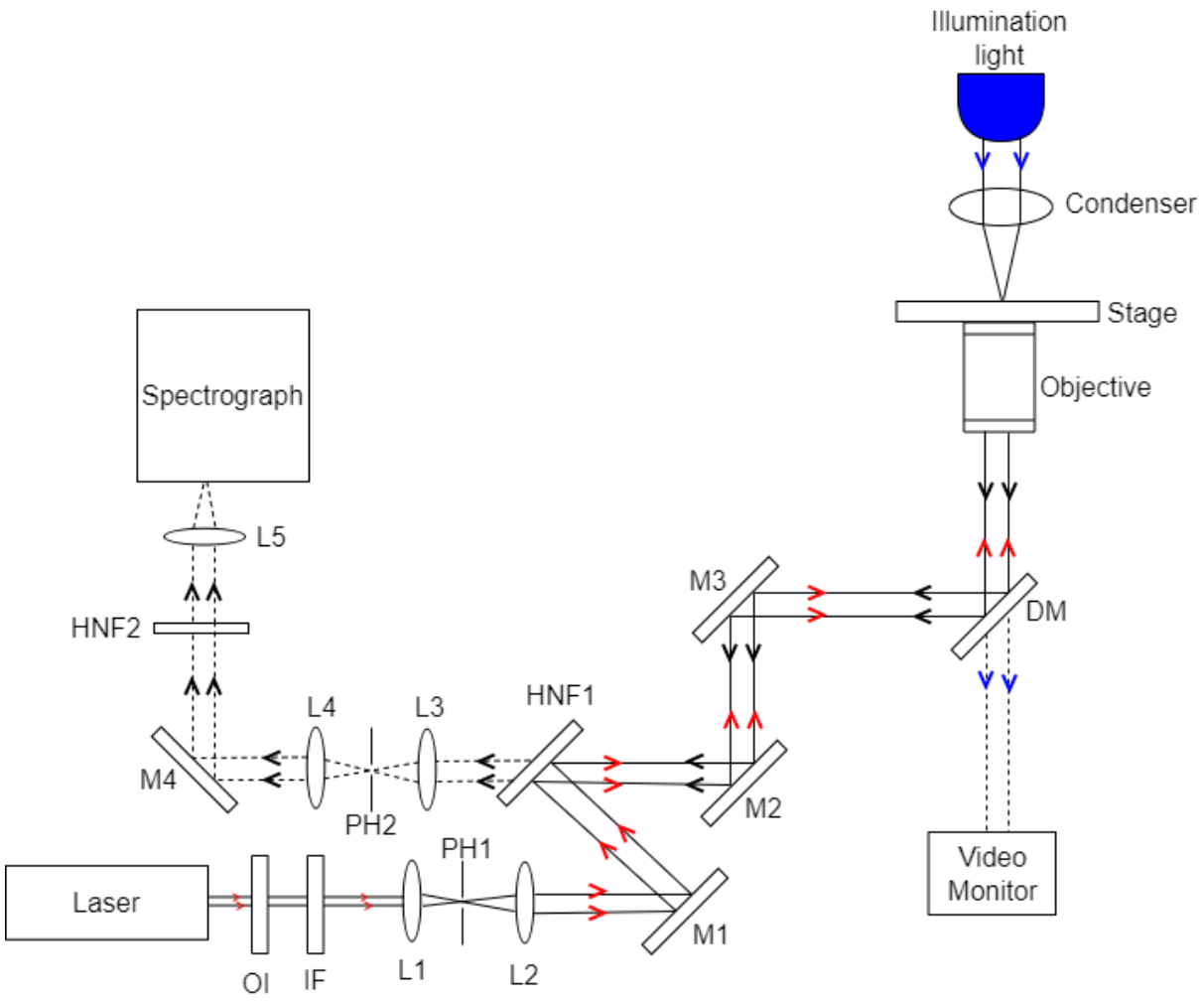

Figure 17. Schematic diagram of the Raman spectroscopy system. The optical components used are optical isolator (OI), interference filter (IF), lens (L1, L2, L3, L4, and L5), pinholes (PH1 and PH2), mirrors (M1, M2, M3, and M4), holographic notch filter (HNF1 and HNF2), dichroic mirror (DM).

This spectrometer contains 3 gratings. The first grating is 1200 grooves/mm with a blazing wavelength at 750 nm, the second is also 1200 grooves/mm and the third is 600 grooves/mm at 750 nm. We used the 600 grooves/mm gratings blazed at 750 nm. PIXIS 100BR CCD camera has a sensor format of 1340x100 pixels, each pixel size of 20x20 μm, which is cooled down to -80 degree C to reduce the thermal noise. With these settings we can record the Raman spectrum from nearly 400 to 2000 cm^{-1} region. The choice of laser wavelength is very important for the Raman

measurement of biological samples especially proteins. Biological samples often exhibit strong fluorescence when excited with visible light source such as blue or green laser. By using a near-infrared (NIR) laser at 780 nm, the fluorescence interference is significantly reduced, which allows a clear Raman spectrum with better signal to noise ratios. Also, near infrared light penetrates the biological samples more effectively than visible light due to the lower absorbance. This reduces the likelihood of the photodamage of the sample.

2.2 Calibration of the Raman System

Each instruments have different responses to photons at different frequencies. It is very important to calibrate the Raman system before measurement of samples. For our system, we used the standard 2 μm polystyrene particles to calibrate for the Raman wavelength shift, whose Raman shifts have been known. The Raman shift for the polystyrene particle is shown in Fig. 18.

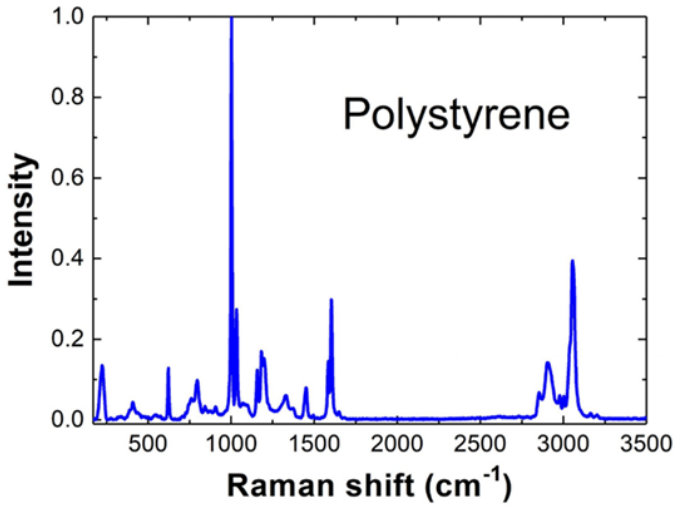

Figure 18. Raman shift of standard polystyrene particle [88].

2.3 Data Analysis

The recorded Raman spectrum raw data was first processed in Origin 2016 software. Deconvolution was also performed in the Origin 2016 software. Then, the PCA, DAPC analysis and Raman barcode was performed in MATLAB.

2.3.1 Data Pre-processing

Preprocessing Raman spectral data is an essential part of the analysis process, designed to improve the quality and clarity of the data before more sophisticated processing techniques are applied. The main objective of preprocessing is to address various artifacts and inconsistencies that may occur during data collection, thereby facilitating more precise and dependable results. Below are some typical preprocessing methods used with Raman spectra, including background subtraction, baseline correction, smoothing, and normalization.

- **Background Subtraction**

The initial spectrum measured includes contributions from the protein samples, the dissolved solvent, and any signals from the optical components used in the measurement process. To isolate the contributions solely from the samples, we first record the spectrum of just the solvent under identical conditions—this is referred to as the background spectrum. We then subtract the smoothed background spectrum from the overall measurement to obtain the spectrum specific to the protein.

- **Baseline Correction**

The measured Raman spectrum typically displays a sloping or curved baseline, potentially caused by fluorescence, Rayleigh scattering, or other optical components such as sample holders.

For accurate spectral comparison, it is crucial to eliminate this baseline. An effective technique is the application of the second derivative algorism to the raw spectra. With this method, any positive peak from the original Raman spectrum translates into the center of a negative peak in the second derivative spectrum. This procedure effectively clears away the sloping or curved backgrounds, resulting in a noticeably sharper negative peak than seen in the original spectrum. In this dissertation work, we used a commercially available software SpectraGryph 1.2 (https://www.effemm2.de/) for baseline correction of all raw spectral data.

- **Smoothing**

One of the main challenges in Raman spectroscopy is the presence of noise within the spectra, which can mask crucial spectral features and hinder the interpretation of data. To counter this, data smoothing is employed as a vital preprocessing step to boost the signal-to-noise ratio (SNR) and enhance the overall quality of the Raman spectra. The fundamental aim of smoothing is to minimize random noise without substantially altering the authentic signal. Sources of noise in Raman spectroscopy can include the laser source, detector electronics, and environmental influences. Smoothing is essential for making the spectral features more prominent and reliable for further analysis, which might include peak identification, quantification, and spectral comparison. Commonly utilized smoothing techniques in Raman spectroscopy encompass the moving average, Savitzky-Golay filter, and Gaussian smoothing. These methods are instrumental in accentuating the main features of the spectrum by averaging out the fluctuations in random noise.

In our application, we utilized a Savitzky-Golay filter with an interval of 5 and a polynomial order of 3, performed in Origin 2016 software. This technique calculates each point in the

smoothed spectrum as a weighted average of the adjacent points from the original spectrum. To maintain symmetry, an equal number of data points is included on each side of the target data points, ensuring an odd total number of points, such as 3, 5, or 7 points in the averages. Fig. 19 shows an example of a raw spectrum, background subtraction, baseline correction, and 5 points Savitzky-Golay smoothing.

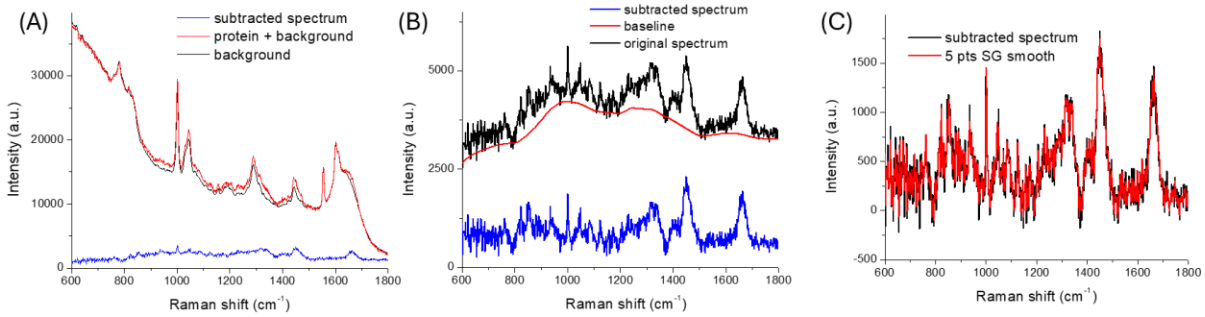

Figure 19. (A) Background subtraction, (B) baseline correction using the second derivative technique, and (C) 5 points Savitzky-Golay smoothing.

Normalization

Before proceeding to further analysis, it's essential to normalize the Raman spectrum to guarantee that the spectra are comparable and standardized. This normalization procedure compensates for variations in experimental conditions, sample concentrations, and instrumental factors variations. The process entails scaling the intensity of the Raman signals either against a standard reference peak or through mathematical techniques that adjust the spectra. A prevalent method for normalization involves setting the total area under the spectrum to 1.0. In this technique, the intensity of each frequency in the spectrum is divided by the square root of the sum of the squares of all intensities. This step ensures consistency across different measurements.

$$I(\nu) = \frac{I(\nu)}{\sqrt{\sum[I(\nu)^2]}} \tag{14}$$

Here, I is the intensity, and *v* is the wavenumbers.

2.3.2 Principal Component Analysis (PCA)

PCA is a dimensionality-reduction method that is used to reduce the dimension of the large data sets by transforming the larger set of variables into a smaller one while preserving the original information. When reducing the number of variables, there might be some loss of original information in the data set, but that outweighs the simplicity of the data presented. This is because the smaller data set is easy to visualize and helps to extract information [89]. We used MATLAB R2020a for the PCA analysis.

Let R be a data matrix of dimension n x p, where p is the number of variables (columns) and n is the number of observations (rows).

$$R = \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1p} \\ r_{21} & r_{22} & \dots & r_{2p} \\ \dots & \dots & \dots & \dots \\ r_{n1} & r_{n2} & \dots & r_{np} \end{bmatrix} \quad (15)$$

The first step to perform the principal component is to standardize the data by subtracting the mean of the data and by dividing it by the standard deviation. Then let X be the normalized matrix.

$$X = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1p} \\ x_{21} & x_{22} & \dots & x_{2p} \\ \dots & \dots & \dots & \dots \\ x_{n1} & x_{n2} & \dots & x_{np} \end{bmatrix} \quad (16)$$

The covariance matrix is the p*p symmetric matrix that contains covariances associated with all possible pairs of initial variables. This covariance matrix summarizes the correlations between all the possible pairs of variables. The covariance matrix of $X_{n\,x\,p}$ is given by

$$C_X = \frac{1}{n-1} X^T X = \begin{bmatrix} \sigma_{11}^2 & \sigma_{12}^2 & \dots & \sigma_{1p}^2 \\ \sigma_{21}^2 & \sigma_{22}^2 & \dots & \sigma_{2p}^2 \\ \dots & \dots & \dots & \dots \\ \sigma_{p1}^2 & \sigma_{p2}^2 & \dots & \sigma_{pp}^2 \end{bmatrix} \quad (17)$$

Here, σ_{ij}^2 is the covariance between the i[th] and j[th] variables.

Next, we calculate the eigenvalue and eigenvector of the covariance matrix. According to the singular value decomposition theorem [90] for a m x n matrix A, whose entries are real or complex numbers, there exists a factorization of the form.

$$A = U\Sigma V^* \quad (18)$$

Where U is a m x m unitary matrix, Σ is a diagonal m x n matrix with diagonal as singular values, V is a n x n unitary matrix, and V^* is the conjugate transpose of V, and it is unitary as well. Hence our data set X can be written as

$$X = U\Sigma V^*$$

$$X^T X = (XU\Sigma V^*)^T (XU\Sigma V^*)$$

$$(X^T X)V = V\Sigma^2 V^* V$$

$$(X^T X)V = V\Sigma^2 \quad (19)$$

Eq. 15 is the eigenvalue problem, where V is the matrix of eigenvectors of $X^T X$ and Σ^2 is the square matrix, whose diagonal elements are the eigenvalues of $X^T X$. Hence

$$C_X V = \frac{1}{n-1}(X^T X)V = \frac{1}{n-1} V\Sigma^2 \quad (20)$$

These eigenvectors, when arranged in decreasing order, form the new dimensions.

Now, we can project the original data on these new rotated axes, and the new data set is given by

$$Z = XV \quad (21)$$

Here, Z is called the score matrix. The columns Z_i's of Z are called Principal Components (PC). They give the variance in the data in decreasing order. Hence, the i^{th} Principal Component is given by

$$PC_i = [x_1 \ x_2 \ x_3 \ldots x_p] \begin{bmatrix} v_{1i} \\ v_{2i} \\ v_{3i} \\ \cdot \\ \cdot \\ \cdot \\ v_{pi} \end{bmatrix} \tag{22}$$

Here, i € [1, n].

The Principal Component is constructed so that the first principal component accounts for the largest possible variance in the data set. That means the first PC is a line that maximizes the variance. The second principal component is uncorrelated with the first PC, which accounts for the next highest variance.

The eigenvectors of the covariance matrix are the directions of the axes where there is the most variance, and these are the principal components. Eigenvalues are the coefficients to the eigenvectors, which give the amount of variance carried in each principal component. By ranking the eigenvector in order of their eigenvalues, highest to lowest, we get the principal component in order of significance. The MATLAB code for the PCA analysis is mentioned in Appendix A.

2.3.3 Discrimination Analysis of the Principal Component (DAPC)

Discriminant analysis of principal components (DAPC) is a statistical method used to identify and describe differences among pre-defined groups based on their measurements. It's

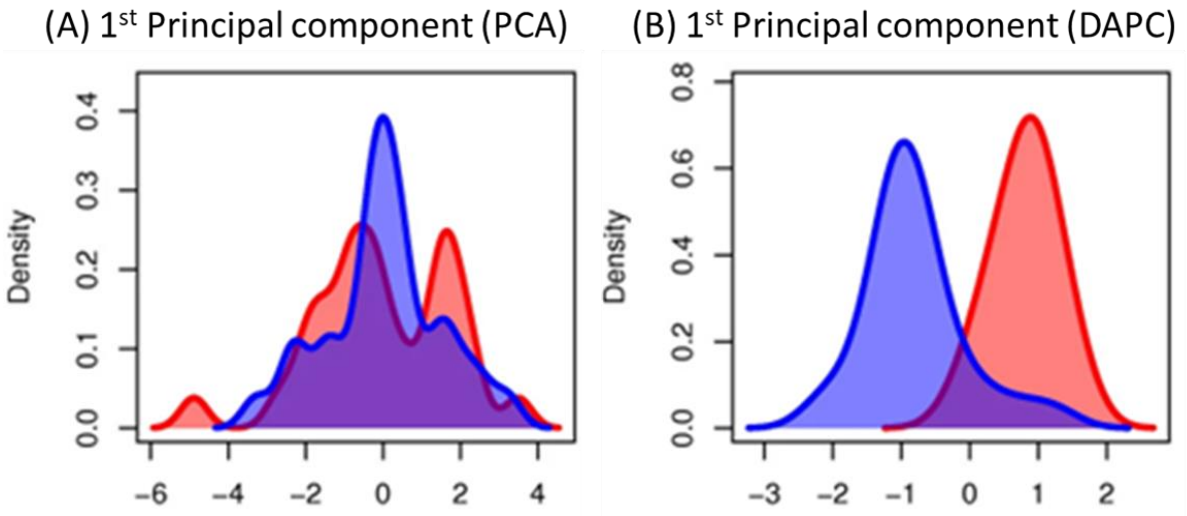

Figure 20. The diagram shows the difference between the PCA and DAPC. (A) shows that the PCA fails to differentiate the groups of data. (B) clearly shows the ability of DAPC to differentiate groups [91].

particularly useful when the dataset is multidimensional, which can often obscure group differences when using traditional methods like principal component analysis (PCA) alone. The discriminant analysis of the principal component uses the k-means clustering, which is the method of vector quantization to partition n observations into k clusters for which each cluster will have the nearest mean. One-way multivariate analysis of variance in MATLAB is used to perform this analysis. This method predicts that the groups in the data set have a similar mean. This technique partitions the data variation between the group and within-group components and maximizes the first one while minimizing the second one [98]. The principal component only looks at the overall variation while the DAPC maximizes the in-between group variation. The MATLAB code for the DAPC analysis is mentioned in Appendix B. The DAPC analysis is a two-step process. The first step involves reducing the dimensionality using the PCA, as described in the previous section. The PCA analysis gives the principal components. The second step involves applying linear discriminant analysis to the principal components. The linear discriminant analysis seeks to find

the axes that best separate the groups in the dataset, maximizing the ratio between group variance and within-group variance.

To compute discriminant analysis of principal components (DAPC), we begin with the matrix Z obtained from principal component analysis (PCA), where each row represents an observation, and each column represents a principal component. Each observation is then assigned to one of k groups, $g_1, g_2, \ldots \ldots g_k$.

Next, we calculate the group mean $\overline{Z}_j$ for each principal component within each group g_j using the formula:

$$\overline{Z}_j = \frac{1}{n_j}\sum_{i=1}^{n_j} z_{ij} \tag{23}$$

Here, $\overline{Z}_j$ is the mean of the j[th] principal component within group g_j, z_{ij} is the value of j[th] principal component for observation i in group g_j, and n_j is the number of observations in group g_j.

After computing the group means, we calculate the within-group matrix W and between group matrix B as follows:

The within group matrix (W) is given by:

$$W = \sum_{j=1}^{k}.\sum_{i=1}^{n_j} n_j(z_{ij} - \overline{Z}_j)(z_{ij} - \overline{Z}_j)^T \tag{24}$$

The between group matrix (B) is given by:

$$B = \sum_{j=1}^{k} n_j(\overline{Z}_j - \overline{Z})(\overline{Z}_j - \overline{Z}))^T \tag{25}$$

Here $\overline{Z}$ represents the overall mean of the principal components.

The total covariance matrix T is then calculated as the sum of W and B:

$$T = W + B \tag{26}$$

We proceed by performing eigen decomposition on the total covariance matrix T to obtain the eigenvalues (λ) and eigenvectors (V) satisfying the equation:

$$TV = \lambda V \tag{27}$$

Next, we select the m eigenvectors corresponding to the m largest eigenvalues to form the matrix V_m. Finally, we transform the original principal component data Z using the selected eigenvectors V_m to obtain the canonical scores (C).

$$C = ZV_m \tag{28}$$

These canonical scores capture the essential discriminatory information between groups and can be plotted to observe differences among groups.

2.3.4 Deconvolution of Raman bands

Peak fitting is the mathematical modeling to adjust the band parameters like amplitude, full width half maximum to obtain the best agreement to the exponential data. Equation (29) is used to fit the experimental data, where $y(v_i)$ is the experimental data as a function of the variable of v, N is the total number of data, and $f(v_i, a)$ is the fitted data.

$$\sum_{i=1}^{N} (y(v_i) - f(v_i, a))^2 \tag{29}$$

Different minimization method is used to minimize the value in Eq. (29), such as the Newton-Raphson method, Gauss-Newton method, Levenberg-Marquard method, and so on. Peak fitting requires an initial guess of the parameters, which will then be the main determinant for the goodness of fit. Those parameters can be each band's width, area, and position. Using this mathematical approach, curve fitting enables the determination of the relative sub volume superposed under the main band. We can use this approach to amide I band of the Raman spectrum of protein, which can be quantified as the protein secondary structures.

2.3.5 Raman barcode

Raman barcode, a novel concept in analytical chemistry and spectroscopy, capitalizes on the distinct Raman spectra exhibited by different substances, like traditional barcodes, for identification and characterization purposes. Unlike conventional barcodes, which rely on patterns of parallel lines of varying thicknesses, Raman barcodes leverage the unique spectral fingerprints of substances obtained through Raman spectroscopy. This technique deconvolutes the overall Raman spectrum into the number of fitting bands. Each fitting band is characterized by three parameters: 1) the position of the band that is the wavenumber, 2) the full width half maximum

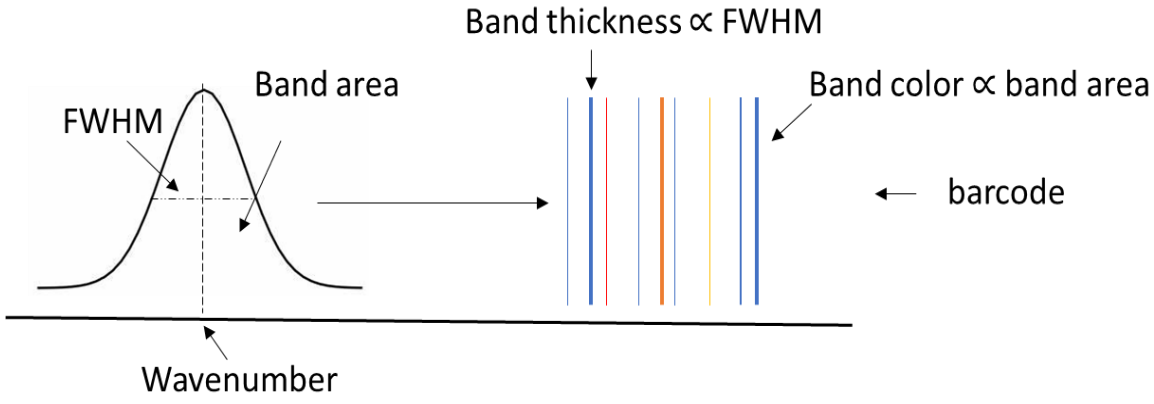

Figure 21. Schematic diagram of Raman bar code generation

(FWHM) of each band, and 3) the total area of the band, as shown in Fig. 21. One line in the Raman barcode corresponds to one fitting band: 1) the position of one barcode line is determined by the wavenumber *v* of the fitting band; 2) the line thickness is chosen to be 1/10 of the FWHM of the fitting band; and 3) the color of one barcode line is proportional to the total area of the fitting band.

2.3.6 Super-resolution Raman imaging

Super-resolution imaging is a set of techniques used in microscopy that allow us to see images at a higher resolution than traditional diffraction-limited resolution with standard optical microscopes. This means that scientists can view objects at the nanoscale, much smaller than what the human eye or regular microscopes can see. These techniques are especially important in fields like biology and materials science, where understanding fine details can help in discovering new things about cell structures, viruses, or new materials. By breaking the limits of normal microscopy, super-resolution imaging helps in providing clearer, more detailed pictures of tiny structures, aiding researchers in making new discoveries and insights.

We utilized ThunderSTORM software to analyze our Raman imaging of R6G molecules on silver nanoparticles. ThunderSTORM is an open-source package tailored for super-resolution microscopy image analysis, particularly for data from techniques like photoactivated localization microscopy (PALM) and stochastic optical reconstruction microscopy (STORM) [92]. ThunderSTORM precisely localizes individual fluorophores within samples, even in densely packed areas, by fitting each detected molecule's point spread function (PSF). This process yields precise coordinates for each molecule, facilitating the reconstruction of super-resolved images beyond the diffraction limit of light microscopy. Consequently, researchers can visualize finer

details and structures within biological samples that were previously inaccessible with conventional microscopy techniques. The software offers a user-friendly interface with a diverse array of analysis options and customizable parameters to accommodate different experimental setups and samples. Users can fine-tune parameters such as localization precision, background estimation, and drift correction to optimize their data analysis. Furthermore, ThunderSTORM provides visualization tools for exploring and analyzing the reconstructed super-resolution images. These tools include image overlays, intensity histograms, scatter plots of molecule positions, and quantitative metrics for evaluating the quality of the reconstruction.

Chapter 3. Characterization and Differentiation of RAS Protein Isoforms by Raman Spectroscopy

3.1 Introduction

Ras proteins, found in all animal cells, are crucial for transmitting signals within cells and belong to a class of proteins called small GTPases [2, 3, 93]. Structurally, the Ras gene consists of two main parts: the G domain and the hypervariable region. The G domain, unique to the Ras superfamily, is essential for cellular processes. Ras proteins function as molecular switches, shifting between inactive guanosine diphosphate (GDP) and active guanosine triphosphate (GTP) states. During this transition, the two-switch region in Ras undergoes conformational changes. The Ras oncogenes, encoded by three genes—K-, H-, and N-Ras—produce proteins of about 188-189 amino acids, sharing 80-90% sequence identity. Variations in the last 19-20 amino acids form the hypervariable region, leading to the creation of different Ras isoforms [2]. Ras mutations are prevalent in roughly 19% of all cancer cases [9, 11, 83]. Notably, mutations in different Ras isoforms are associated with different types of cancer [11, 94]. Understanding these distinctions is crucial for unraveling the different roles of Ras isoforms.

Until recently, the RAS oncogene was deemed impervious to pharmacological intervention. However, recent breakthroughs have revealed the druggability of the KRAS G12C protein through the development of small inhibitor molecules. Notably, the emergence of ARS-1620, a novel inhibitor explicitly targeting the mutant KRAS G12C, presents promising prospects for targeted therapeutic interventions in cancer [83]. This discovery signifies a pivotal advancement in the realm of RAS-targeted cancer therapy. Efforts towards developing such targeted therapeutics

necessitate a profound understanding of the distinctive properties inherent to Ras isoforms, alongside expeditious methods for their identification and differentiation.

Raman spectroscopy, a spectroscopic technique commonly utilized for determining the vibrational modes of molecules, serves as a valuable tool for providing a structural fingerprint essential for molecule identification [95, 96]. It has been widely used to study biological samples, including human cells [10] and proteins [97, 98]. However, due to the substantial sequence similarity among Ras isoforms, distinguishing their Raman spectra requires further data analysis. Multivariate analysis techniques, such as principal component analysis and discrimination analysis of principal components, prove indispensable for discerning highly similar Raman spectra. This analytical approach finds applicability across various sample types, ranging from human urine [99, 100], normal and cancerous tissues [101, 102], to protein and protein-drug complexes [88].

Harda et al. [103] investigated protein tyrosyl residues through the characteristic tyrosine doublet in the Raman spectrum, with a specific focus on the vibrations of the benzene ring as an indicator of tyrosine's hydrogen bonding within the side chain. The intensity ratio of tyrosine's doublet at 855 cm^{-1} to 830 cm^{-1} indicates its role in hydrogen bonding being 1.25 as neutral, with higher ratios suggesting a greater propensity to accept hydrogen bonds. Conversely, lower ratios point to a stronger hydrogen bond donation. The intensity ratio of the Fermi doublet peaks around 1240 cm^{-1} to 1220 cm^{-1} also sheds light on the hydrophobic or hydrophilic nature of the phenol side chain in the surrounding environment. The ratio above 1.1 represents the hydrophobic solvent, whereas below 0.9 represents the hydrophilic solvent. Deconvolution of the Amide I band has been standard in estimating the secondary structure of various proteins [79, 104]. Additionally, according to Fontana et al [18], structural changes in protein conformations in solutions have been

assessed through PCA. The interaction between proteins and drugs, such as the anticancer protein Azurin and p53, was explored by Signorelli et al through changes in the microenvironment observed in the tyrosine and tryptophan doublets [88]. Additionally, changes in the complex's secondary structure were observed by fitting the amide I band. Wang et al. utilized 18O isotope-edited Raman difference spectroscopy in combination with ab initio calculations to study the phosphate modes for GDP and GTP binding to RAS [80]. For the classification of colon cancer cells with mutations in the kras gene, Liu et al. [23] employed laser tweezer Raman spectroscopy, along with PCA and linear discrimination analysis (PCA-LDA). Raman barcode is the recently proposed technique that can be generated using the deconvoluted peaks. This will be helpful for the quick identification and record-keeping of the Ras isoforms [105]. The advancement of technology such as AI holds promise for the practical use of Raman barcodes generated from deconvoluted peaks for quick identification of Ras isoforms. Our literature review did not yield studies characterizing Ras isoforms outside of cells with Raman spectroscopy, signifying an untapped area of research.

In this work, we differentiated the different Ras isoforms by combining the normal Raman spectroscopy and DAPC model. We analyzed the amide I band, the intensity ratio of tyrosine, and the Fermi doublet to characterize Ras isoforms based on the deconvoluted Raman peaks. We also generated the Raman barcode for different Ras isoforms. Finally, we utilized Raman spectroscopy to observe the effect of ARS-1620 inhibitor molecules on the KRAS G12C protein isoform.

The list of Ras isoforms measured using the normal Raman spectroscopy is shown in Table 2. The primary amino acid sequence for the studied Ras isoforms is given in the Table 3.

Table 2. List of different Ras and their isoforms measured in this experiment.

Sample Name	Nucleotide	Concentration(mg/ml)
GG Hs KRAS4a(2-189)	GDP	9.35
GG Hs KRAS4b (2-188)	GDP	28.44
GG Hs KRAS4b (2-188) G12C	GDP	10.30
G Hs HRAS(1-189)	GDP	17.49
G Hs NRAS(1-189)	GDP	12.26
G Hs KRAS4b(1-169)G12C	GDP	25.2
G Hs KRAS4b(1-169)G12C/ARS-1620 1st	GDP	14.4
G Hs KRAS4b(1-169)G12C/ARS-1620 2nd	GDP	13.3
G Hs KRAS4b(1-169)G12C/ARS-1620 3rd	GDP	11.6
GG Hs KRAS4a (2-189)	GppNHp	9.6
GG Hs KRAS4b (2-188)	GppNHp	19.7
GG Hs KRAS4b (2-188)G12C	GppNHp	16.4
G Hs HRAS(1-189)	GppNHp	9.6
G Hs NRAS(1-189)	GppNHp	16.1

Table 3. The primary amino acid sequence of Ras isoforms

GG-Hs.KRAS4a(2-189)	GGTEYKLVVVGAGGVGKSALTIQLIQNHFVDEYDPTIEDSYRKQVVIDGETCLLDILDTAGQEE YSAMRDQYMRTGEGFLCVFAINNTKSFEDIHHYREQIKRVKDSEDVPMVLVGNKCDLPSRTVD TKQAQDLARSYGIPFIETSAKTRQRVEDAFYTLVREIRQYRLKKISKEEKTPGCVKIKKCIIM
GG-Hs.KRAS4b (2-188)	GGTEYKLVVVGAGGVGKSALTIQLIQNHFVDEYDPTIEDSYRKQVVIDGETCLLDILDTAGQEE YSAMRDQYMRTGEGFLCVFAINNTKSFEDIHHYREQIKRVKDSEDVPMVLVGNKCDLPSRTVD TKQAQDLARSYGIPFIETSAKTRQGVDDAFYTLVREIRKHKEKMSKDGKKKKKKSKTKCVIM
GG-Hs.KRAS4b (2-188) G12C	GGTEYKLVVVGACGVGKSALTIQLIQNHFVDEYDPTIEDSYRKQVVIDGETCLLDILDTAGQEE YSAMRDQYMRTGEGFLCVFAINNTKSFEDIHHYREQIKRVKDSEDVPMVLVGNKCDLPSRTVD TKQAQDLARSYGIPFIETSAKTRQGVDDAFYTLVREIRKHKEKMSKDGKKKKKKSKTKCVIM
G-Hs.HRAS(1-189)	GMTEYKLVVVGAGGVGKSALTIQLIQNHFVDEYDPTIEDSYRKQVVIDGETCLLDILDTAGQEE YSAMRDQYMRTGEGFLCVFAINNTKSFEDIHQYREQIKRVKDSDDVPMVLVGNKCDLAARTVE SRQAQDLARSYGIPYIETSAKTRQGVEDAFYTLVREIRQHKLRKLNPPDESGPGCMSCKCVLS
G-Hs.NRAS(1-189)	GMTEYKLVVVGAGGVGKSALTIQLIQNHFVDEYDPTIEDSYRKQVVIDGETCLLDILDTAGQEE YSAMRDQYMRTGEGFLCVFAINNSKSFADINLYREQIKRVKDSDDVPMVLVGNKCDLPTRTVD TKQAHELAKSYGIPFIETSAKTRQGVEDAFYTLVREIRQYRMKKLNSSDDGTQGCMGLPCVVM
G-Hs.KRAS4a(1-169) G12C	GMTEYKLVVVGACGVGKSALTIQLIQNHFVDEYDPTIEDSYRKQVVIDGETCLLDILDTAGQEE YSAMRDQYMRTGEGFLCVFAINNTKSFEDIHHYREQIKRVKDSEDVPMVLVGNKCDLPSRTVD TKQAQDLARSYGIPFIETSAKTRQGVDDAFYTLVREIRKHKEK

3.2 Materials and Methods

3.2.1 Raman Measurements

The normal Raman measurements were performed using near-infrared Raman micro-spectroscopy. The excitation source for the measurements was 780 nm laser (TEC-300–0780–0500, Sacher Lasertecknik Inc.). The scattered Raman signal was collected using the 100X oil objective having NA of 1.3 and recorded with Triax 320 imaging spectrograph. For the Raman measurements of Ras proteins, ~ 5 uL liquid protein sample was taken on a cleaned UV silica plate mounted on a sample holder, covered with a glass coverslip and made airtight using vacuum grease to avoid the vaporization of the liquid protein inside the sample holder. Then, the sample holder containing protein was transferred to the microscope to record the Raman spectra. We also measured the matched buffer under identical conditions. For all measurements, the laser power was 20 mW, and the integration time was 60s with 10 accumulations. The use of high energy laser might have some effect on the protein samples such as photochemical or photothermal damage. To avoid these effects, we slightly shifted the Raman measurement position during each acquisition.

3.2.2 Data Analysis

We used different software to analyze the data. First, we transferred the data to the Origin Pro 2016 software. We used the finger-printing 600 cm^{-1} – 1800 cm^{-1} region of the collected data for further data processing. Since different proteins were in different buffer solutions, we recorded and subtracted the corresponding buffer's Raman spectrum from each protein's spectrum. After the background subtraction, the data was smoothed, and the baseline was corrected for each spectrum using SpectraGryph 1.2 (https://www.effemm2.de/). All the spectra were then

normalized to the peak value at 1003 cm^{-1} for comparisons. The Amide I band of protein was fitted using peak fit functions in origin pro-2016. The PCA and DAPC were performed using MATLAB.

3.3 Results and Discussions:

3.3.1 Raman Spectra Can Differentiate Ras Isoforms

Normal Raman spectra for studied GDP and GppNHp isoforms are measured in buffer solutions using near-infrared Raman spectroscopy. Under the same condition, the buffer spectrum was also measured and subtracted from the protein spectrum to correct for the solvent background. Then, the spectrum was smoothed, and the baseline was corrected. Since the 1003 cm^{-1} Phe protein band was less dependent on the solvent property, we normalized all spectra to this band for comparison. This way, we created the Raman spectrum library for RAS isoforms for GDP and GppNHp, as shown in Fig. 22 (A) and (B), respectively. This also serves the Ras isoform's Raman spectrum library. We repeated the sample measurements three times and each time we measured 30 spectra to find out the measurement reproducibility. The plotted spectra in Fig. 22 are the average of 90 spectra. Since all isoforms have a high level of sequence similarity, it is very difficult to differentiate between isoforms without further data processing of the spectra. To discriminate all isoforms from one another, we created the model based on the principal component analysis (PCA) and discrimination analysis of the principal component. PCA transforms a set of correlated variables into a set of uncorrelated variables called principal components. Discrimination analysis of the principal component combines PCA and discriminant analysis (DA) to identify and describe clusters of isoforms.

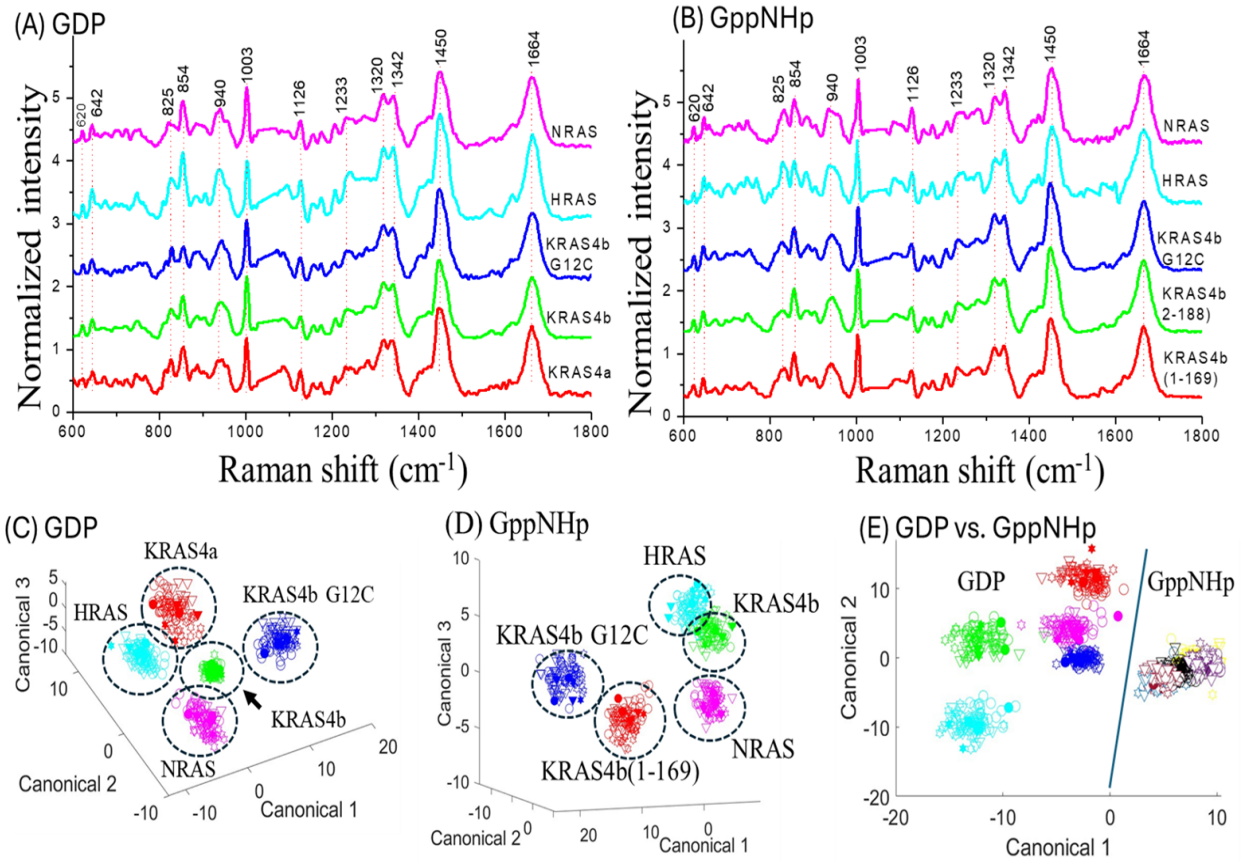

Figure 22. (A) Normal Raman spectra of GDP isoforms of KRAS4a, KRAS4b, KRAS4b G12C, HRAS and NRAS. (B) Raman spectra of GppNHp isoforms for KRAS4b(1-169), KRAS4b(2-188), KRAS4b G12C, HRAS and NRAS. All the measurements were performed using 30mw laser power and the acquisition time was 40s with 10 accumulations. Buffer background under identical condition was measured and subtracted for background correction. The background subtracted spectra were smoothed, baseline corrected and normalized to the peak intensity at 1003 cm^{-1}. Each spectrum is the average of 3 independent measurements containing 30 spectra for all 5 protein isoforms. (C) DAPC plot showing the discrimination within the GDP isoforms (D) DAPC plot showing the discrimination within the GppNHp isoforms and (E) DAPC plot showing the discrimination between the GDP and GppNHp isoforms. Each DAPC plot has 3 different shapes (circular, star and triangular) with two types (hollow and filled). The three different shapes represent the 3 independent measurement for the same isoform and hollow data is for training data and filled data for testing isoforms.

In this work, we first separated each measurement into training and testing spectra. We started with GDP isoforms. Out of 30 spectra measured in one experiment, we used 28 spectra to train a model and the remaining 2 spectra to test that model. Then, we combined all training data sets for all 5 isoforms in GDP. PCA analysis was performed on the combined training data set,

and the principal component coefficient was used to perform the DAPC analysis. After this analysis, the first three canonical variables were plotted, as shown in Fig. 22(C). We then projected the testing data to the same plot to see whether these data go to the right cluster or not. Interestingly, we observed all testing spectra went to the right cluster. In Fig. 22(C), each color represents the one isoform. In each color, there are 3 different shapes: circular, triangular and star. Three shape represents the triplicate measurement of same isoform to test the reproducibility of measurements. Also, in each shape they are filled and unfilled. Filled shape are for testing data and unfilled shape are for training data. From Fig. 22 (C), we observed that each cluster are separated from each other, indicating that we can differentiate between the GDP isoforms.

We repeated the same process for GppNHp isoforms and plotted them in Fig. 22(D). The result shows that we can also differentiate the GppNHp isoforms by DAPC analysis. In particular, a stronger contribution from KRAS4b G12C than NRAS was observed in the Canonical 1 spectrum, since the NRAS cluster is aligned with the negative side of the Canonical 1 axis while KRAS4b G12C aligns on the positive side. When analyzing the original spectra, we observed that the lipid band was higher for KRAS4b G12C. Also, the 830 cm^{-1} band is higher, and the 855 cm^{-1} band is wider for KRASG12C than NRAS.

Next, we extended this technique to differentiate between GDP and GppNHp isoforms. We combined all 10 isoform's training data, plotted the results, and projected the testing data to the same canonical space. We observed a clear separation between GDP and GppNHp isoforms, as shown in Fig. 22(E).

The grouping of all three shapes within the same cluster suggests the reproducibility of the measurements. Within each cluster, there are some variations of the points in canonical space, which can be due to the noise in the spectra or the measurement uncertainty of the Raman system.

When projecting the testing data, we observed that all filled shapes went to the right cluster, showing the usefulness of this technique in identifying the unknown spectra among this isoform. The tight cluster represents less noise in the spectrum and, in this case, due to the higher sample concentration.

3.3.2 Deconvolution of Raman Spectra and Raman Barcode Differentiate Ras Isoform

Deconvolution separates the overlapping Raman bands into their individual components by fitting a mathematical model to the spectrum. The deconvoluted band estimates the peak position,

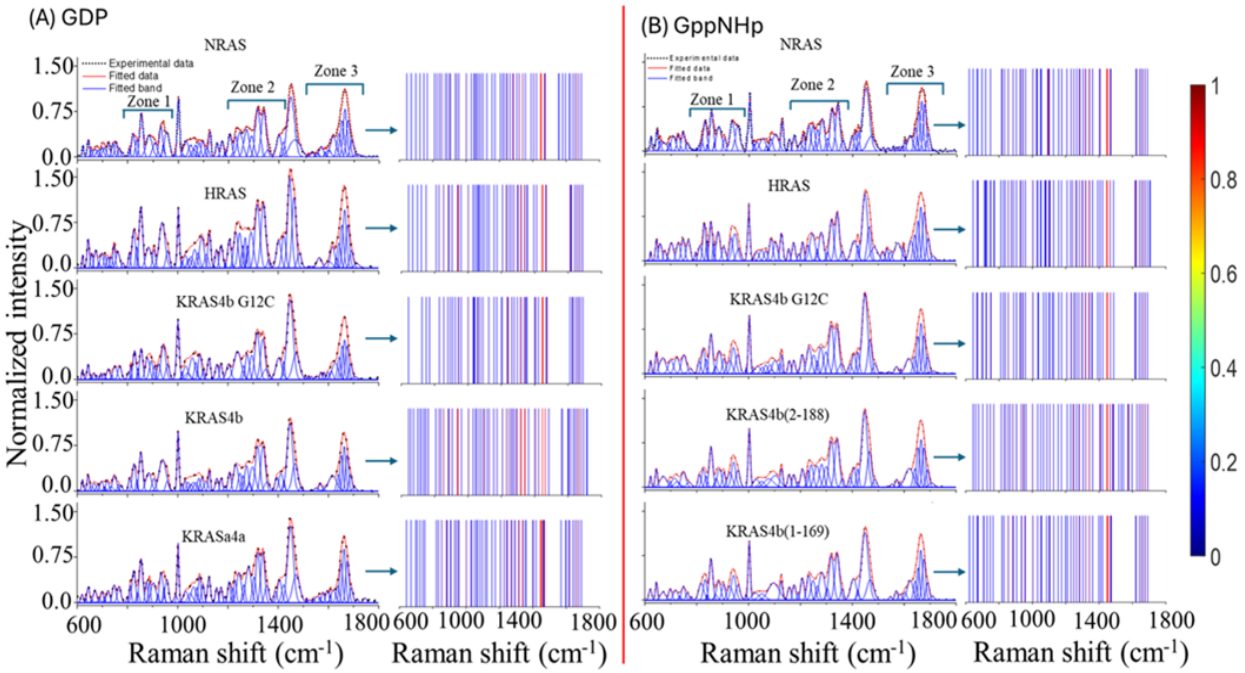

Figure 23. Deconvolution of Raman spectra and Raman barcodes of Ras isoforms. (A) The deconvoluted Raman bands (left) and Raman barcodes (right) for GDP isoforms. (B) The deconvoluted Raman bands (left) and Raman barcodes (right) for GppNHp isoforms. The dotted line represents the original Raman spectra data, the red line is the overall fitting data and blue bands inside represent the fitted bands. The Raman barcode was generated using the percentage of area covered by each band and FWHM of each band. The fitting data in three zones (zone 1, zone 2, and zone 3) were used to calculate tyrosine doublet ratio, Fermi doublet ratio and protein secondary structure.

width, and intensity of each peak. All the GDP and GppNHp isoforms were deconvoluted and shown in Fig. 23. The three zones mentioned in the deconvoluted spectrum were used to characterize the isoforms further.

The Raman barcodes were generated using the deconvoluted band information. Any deconvoluted bands that contribute less than 5% towards the total band area of the original Raman spectrum was removed for barcode generation. Our algorithm converts the band sequence to a barcode, by 1) assigning each fitting band to a line whose thickness is 1/10 of the width of that fitting band; 2) plotting the line on the corresponding wavenumber; and 3) assigning the line color to represent the area of each fitting band. With the Raman barcodes, we can easily see the variation among Ras isoforms, which we cannot be directly observed from spectra. Therefore, the Raman barcode serves as the electronic record keeping for these isoforms, which can be easily converted back to the original spectrum or bands. Also, with the future advancements of artificial technology and by using appropriate computer software or mobile applications there will be the possibility to identify or cross check the isoforms.

3.3.3 Spectral Characterizations of Amide Band, Tyrosine Doublet, and Fermi Doublet Band of Ras Isoforms

Fig. 24 shows spectral characterizations of Ras isoforms. Fig. 24(A) displays the typical deconvoluted Raman bands, focusing on three zones to characterize the properties of Ras isoforms. Zone 1 includes the Tyrosine doublet at 830 cm^{-1} and 853 cm^{-1}, which indicates hydrogen bonding interactions involving phenolic OH rings and the strength and nature of these bonds. We maintained the same FWHM when fitting the bands to compare doublet amplitudes consistently. Fig. 24C(I) shows that HRAS and NRAS isoforms have higher Tyr doublet ratios (I_{853}/I_{830}) in

GDP, suggesting a greater tendency for hydrogen bond acceptance. However, Fig. 24D(I) shows smaller Tyr doublet ratios for these isoforms in GppNHp, reflecting a change in hydrogen bonding conditions between GDP and GppNHp.

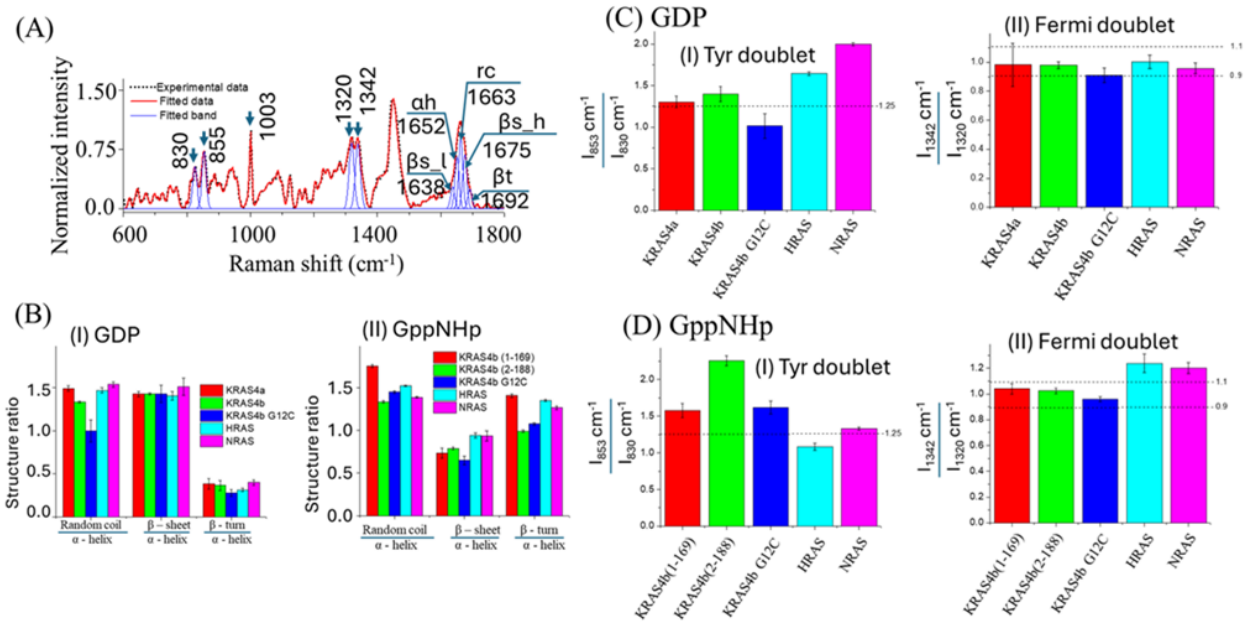

Figure 24. Spectral characterizations of Ras isoforms. (A) shows a representative Raman spectrum with deconvoluted bands for Tyrosine doublets at 830 cm^{-1} and 855 cm^{-1}, a Fermi doublet at 1320 cm^{-1} and 1340 cm^{-1}, and the Amide I band at 1665 cm^{-1}, which is split into five bands. The 1638 cm^{-1} band corresponds to the lower frequency beta-sheet (βs_l), 1652 cm^{-1} to the alpha helix (αh), 1663 cm^{-1} to the random coil (RC), 1675 cm^{-1} to the beta-turn (βt), and 1692 cm^{-1} to the high-frequency beta-sheet (βs_h). (B) presents the ratios of the beta-sheet, random coil, and beta-turn to alpha helix for (I) GDP and (II) GppNHp. (C) and (D) display bar diagrams showing the ratios of tyrosine and Fermi doublets for GDP and GppNHp isoforms, respectively.

Zone 2 features the Fermi doublet peaks at 1342 cm^{-1} and 1320 cm^{-1}, arising from Fermi resonance—an anharmonic coupling between vibrational modes causing the energy levels to interact and split, forming a doublet. The intensity ratio between these two peaks (I_{1342}/I_{1320}) provides insights into the protein's hydrophobicity; a higher Fermi doublet ratio suggests more internal hydrophobic residues, whereas a lower Fermi doublet ratio indicates greater exposure on the protein surface. According to the data in Fig. 24C(II), GDP isoforms have a neutral

environment with ratios between 0.9 and 1.1. For GppNHp isoforms, Fig. 24D(II) show an increased hydrophobicity in HRAS and NRAS, indicating more buried hydrophobic structures.

Zone 3 contains the amide I band around 1665 $cm^{-1,}$ which arises from the C=O stretching vibrations of the peptide backbone, which are sensitive to the protein's secondary structure. From the deconvoluted band position and area of each band, we can predict the percentage of secondary structures. The secondary structures are fitted with beta-sheet (1630 – 1650 cm^{-1} for lower frequency and 1670 -1680 cm^{-1} for higher frequency), alpha helix (1650 -1660 cm^{-1}), random coil (1660 – 1670 cm^{-1}) and beta-turn (1680 -1700 cm^{-1}). From the fitting data, we calculate the intensity ratios of these bands (including random coil/α-helix, β-sheet/α-helix, and β-turn/α-helix). For GDP isoforms, we observed that the dominant structures are the random coil and beta-sheet, and the beta-turn amount is the lowest, compared to the alpha-helix. For GppNHp, we observed that the ratio for the beta-sheet is the lowest.

3.3.4 Characterization of KRAS G12C and ARS-1620 interaction

Fig. 25 shows characterization of the interaction between KRAS G12C protein and ARS-1620 drug using Raman spectroscopy. Fig. 25(A) shows the Raman spectra of KRAS G12C protein and KRAS G12C/ ARS-1620 complex. Three independently prepared complexes were measured and plotted as complex1, complex2, and complex3. The shaded regions highlight the change in the complex's spectrum after adding the ARS-1620 drug in Tyrosine doublet (830 - 853 cm^{-1}), Fermi doublet (1342 - 1320 cm^{-1}), and the amide I band (~ 1665 cm^{-1}). From Fig. 25(B), we can observe the change in the protein's structure after adding the ARS-1620 drug. Alpha-helix structure is considered the most stable structure in the protein. The bar diagram in Fig.25(B) shows that the ratios of the random coil and beta-turn to alpha helix was increased for the complex,

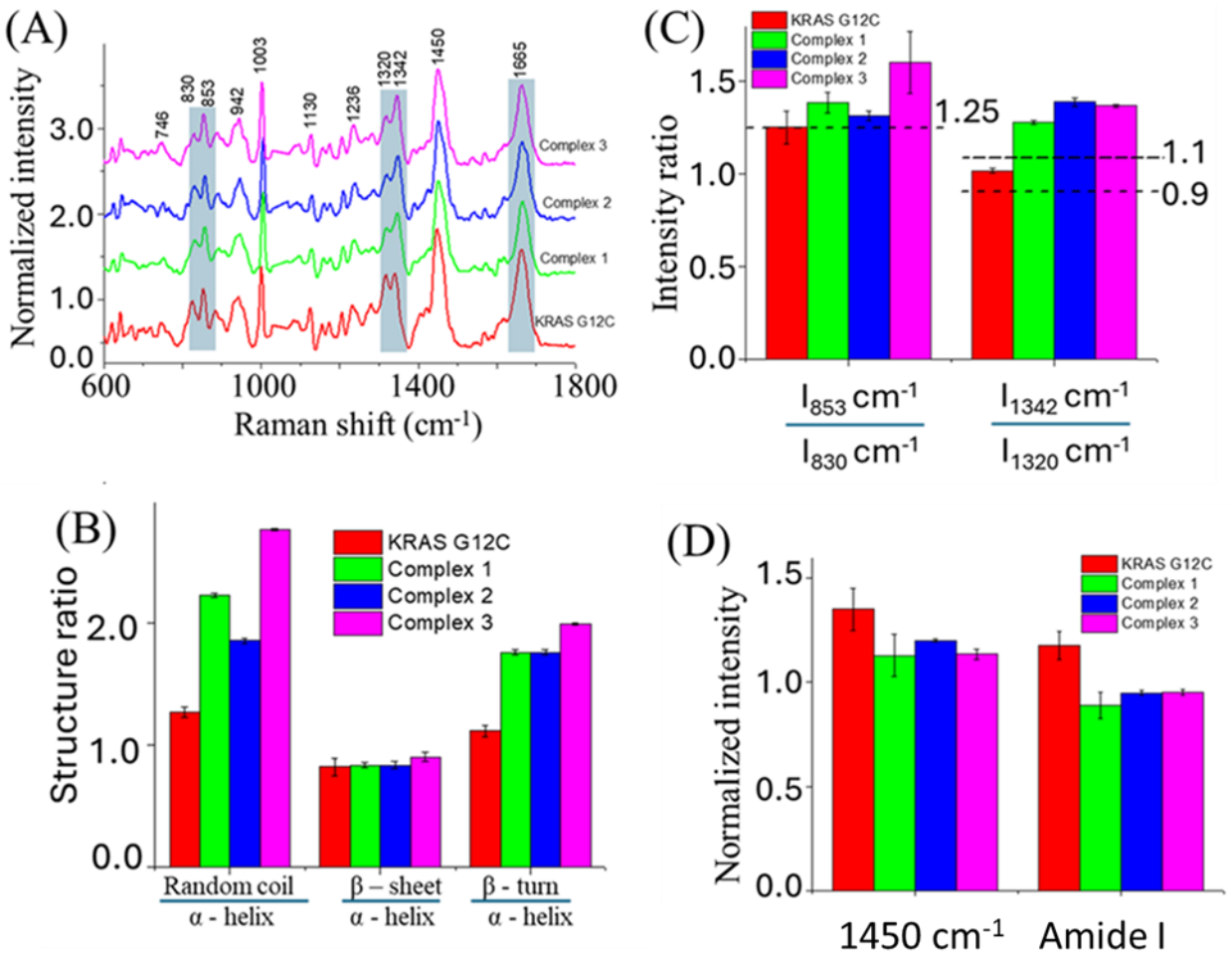

Figure 25. Characterization of the interaction between KRAS G12C protein and ARS-1620 drug. (A) Raman spectra of the KRAS G12C protein and the KRAS G12C/ARS-1620 drug complexes. All the spectra were smoothed, baseline corrected and normalized to 1003 cm^{-1} band. The laser power was 30 mA and acquisition time was 40 s and accumulatio was 10 times. The shaded part shows the changes in the Raman spectrum after the interaction with drug molecules. (B) secondary structures ratio for protein and protein drug complex. (C) bar diagram showing the tyrosine and Fermi doublet intensity ratio. (D) bar diagram showing the overall decreases in the 1450 cm^{-1} band and Amide I band at the 1665 cm^{-1} after the interaction with ARS-1620.

Fig.25(C) shows the tyrosine doublet and Fermi doublet intensity ratios. Fig.25(D) plots the bar diagram showing the overall decreases in the 1450 cm^{-1} band and Amide I band at the 1665 cm^{-1} after the interaction with ARS-1620 suggesting the increase in the percentage of random coil and beta-turns structures in protein's amide I band. This finding suggests that the KRAS G12C protein became more unstable after the addition of the ARS-1620 drug. In addition, the tyrosine

doublet and Fermi doublet ratios were calculated and shown in Fig. 25(C). The tyrosine ratio (I_{853}/I_{830}) was increased after the addition of the drug, indicating the strength of the hydrogen bond increased. Similarly, the Fermi doublet ratio (I_{1342}/I_{1320}) was also increased after the addition of the drug, indicating the increased hydrophobicity with the drug's addition. This means the protein has buried its hydrophobic residue from the protein surface after adding drugs. The 1450 cm^{-1} lipid band is due to the CH_2 bending vibration of aliphatic amino acids. The aliphatic amino acids contain side chains with methylene (CH_2) groups, and these contribute to the vibrational modes and can be detected in the Raman spectrum. So, this peak is due to the bending motion of CH_2 groups in the amino acid residues. The decrease in the intensity of this band (I_{1450})) is, therefore, associated with the unfolding of the protein structures after interacting with ARS-1620.

3.3.5 Conclusions

We have successfully built a Raman spectrum library for Ras protein isoforms. Using normal Raman spectroscopy and discriminant analysis of principal components (DAPC), we differentiated Ras isoforms within GDP and GppNHp states, as well as between these states. GppNHp isoforms align on the positive canonical 1 axis, while GDP isoforms align on the negative canonical 1 axis, indicating a distinct separation between GDP and GppNHp isoforms in canonical space. The DAPC model can accurately categorize the testing set, which validate its effectiveness in identifying unknown protein spectra. The deconvoluted bands have facilitated the creation of a Raman barcode, aiding future isoform identification through AI-powered software. The HRAS and NRAS exhibit a higher tyrosine ratio when GDP-bound, indicating stronger hydrogen bonding. The Fermi doublet ratio is also higher for GDP-bound HRAS and NRAS, suggesting residues are more buried in the hydrophobic core. Upon ARS-1620 interaction, we observed a

reduction in KRAS G12C secondary structures, increased the strengths of tyrosine's hydrogen bonding, and increased hydrophobicity, with the decreased 1450 cm^{-1} band intensity suggesting protein unfolding.

Chapter 4. Enhancement of Raman Spectra by Vacuum Evaporation

4.1 Introduction

Raman spectroscopy is a powerful analytical technique for identifying and detecting different molecules and has been widely applied in biology [106], medicine [107], nanomaterials [108], food science [109], and in the study of gases [110], chemicals compounds [111] and so on. Raman spectroscopy is suitable for non-destructive label-free analysis and can give qualitative as well as quantitative data of the analyte molecules [112]. Raman spectra of the analyte molecules arise from inelastic scattering when incident photons from a laser source at a given frequency impinge upon the sample [113]. However, the molecular cross-section of Raman scattering is normally very low so that only a small number of scattered photons are generated as the Raman signal that carry on the vibrational frequency of the analyte molecules. This weak signal greatly limits the applications of the Raman technique for chemical and biological detection of molecules, especially for low-concentration samples. Therefore, the enhancement of the weak signal is necessary for broad applications. There have been many enhancement techniques for increasing Raman signals, including resonance Raman spectroscopy [21, 22], surface enhancement Raman scattering (SERS) [15, 16], tip enhancement Raman scattering (TERS) [17, 26], and fibre enhancement Raman scattering (FERS) [19, 20]. Resonance Raman spectroscopy enhances the Raman signal when the excitation wavelength is close to or overlapping with the absorption or emission wavelength of molecules [81]. The resonance Raman technique requires the tuneable laser source or the laser source whose frequency matches the

transition frequencies of the molecules to get the enhancement [114]. SERS uses metallic nanoparticles to enhance the Raman signal when the analyte molecules are adsorbed on the surfaces of metal nanoparticles due to the combination of electromagnetic and charge transfer effects [115]. The signal enhancement largely depends on the shape and size of the nanoparticles [116]. Although the SERS with nanoparticles has the strongest enhancement over normal Raman spectroscopy, it has poor reproducibility and generally produces the alteration of Raman spectra because of the difficulty in controlling the nanoparticle's size/shape and the location of the adsorbed analyte molecules on the surface [117]. The TERS technique combines scanning probe microscopy and surface-enhanced Raman spectroscopy [117] with a nano-sized metal tip or metal nanoparticles on the tip. One of the shortcomings of TERS is that the tip can be easily erased [17]. Fiber enhancement Raman spectroscopy technique (FERS) has been applied to enhance the Raman signals of gases and antibiotic molecules filled in a long hollow fiber [36, 118]. In FERS, the analyte solution or gas is filled in a long hollow optic fiber in which the excitation light and Raman scattering light are propagating through so that the effective interaction volume between the excitation light and analyte molecules can be increased and hence, a higher Raman signal can be obtained [36]. Since the analyte molecules are filled in a small hollow vessel, changing the sample, and cleaning the vessel are relatively difficult.

The working principle of vacuum-enhanced micro-Raman spectroscopy (VERS) relies on the increase in local concentration of analyte molecules in the excitation volume of micron-sized laser focus by the assistance of vacuum evaporation of solvents. In this method, the analyte molecules dispersed in a solvent solution with a desired volume are kept in a sealed sample holder which is connected to the vacuum pump and placed on a

micro-Raman microscope stage. As the vacuum pump is turned on, the solvent liquid inside the sample holder starts evaporating, leading to an increase in the concentration of analyte molecules such that more analyte molecules are scattered by the excitation light in the effective volume of the laser focus, and hence produce higher Raman signal.

The advantages of the VERS technique include: (1) it does not require any sensor materials such as metal nanoparticles to enhance the Raman scattering signal as required in SERS and TERS, and hence it generates the same Raman spectra as in the normal Raman spectroscopy; (2) it is much easier and simpler to change different samples and clean the sample holder compared to FERS. Therefore, VERS is user-friendly and robust for analysing different molecules, biomolecules, and antibiotics with a variety of solvents; (3) the VERS technique is particularly useful for detecting analyte molecules dissolved in some solvents (such as DMSO) that have a strong Raman background and usually interfere the weak Raman signals of analyte molecules. As the result of vacuum evaporation of the solvent, the effect of solvent's Raman background can be significantly reduced such that the weak Raman signals of analyte molecules can be detected; (4) the VERS technique could be combined with gold or silver nanoparticles to further increase the sensitivity of SERS.

4.2 Materials and Methods

- **Vacuum-Enhanced Micro-Raman Spectroscopy**. The schematic of vacuum enhancement micro-Raman spectroscopy is shown in Fig. 26(a, b). Briefly, the excitation laser beam is tightly focused with a microscope objective to form a micron-sized focus and Raman scattered light of analyte molecules from the focus region is collected with the same

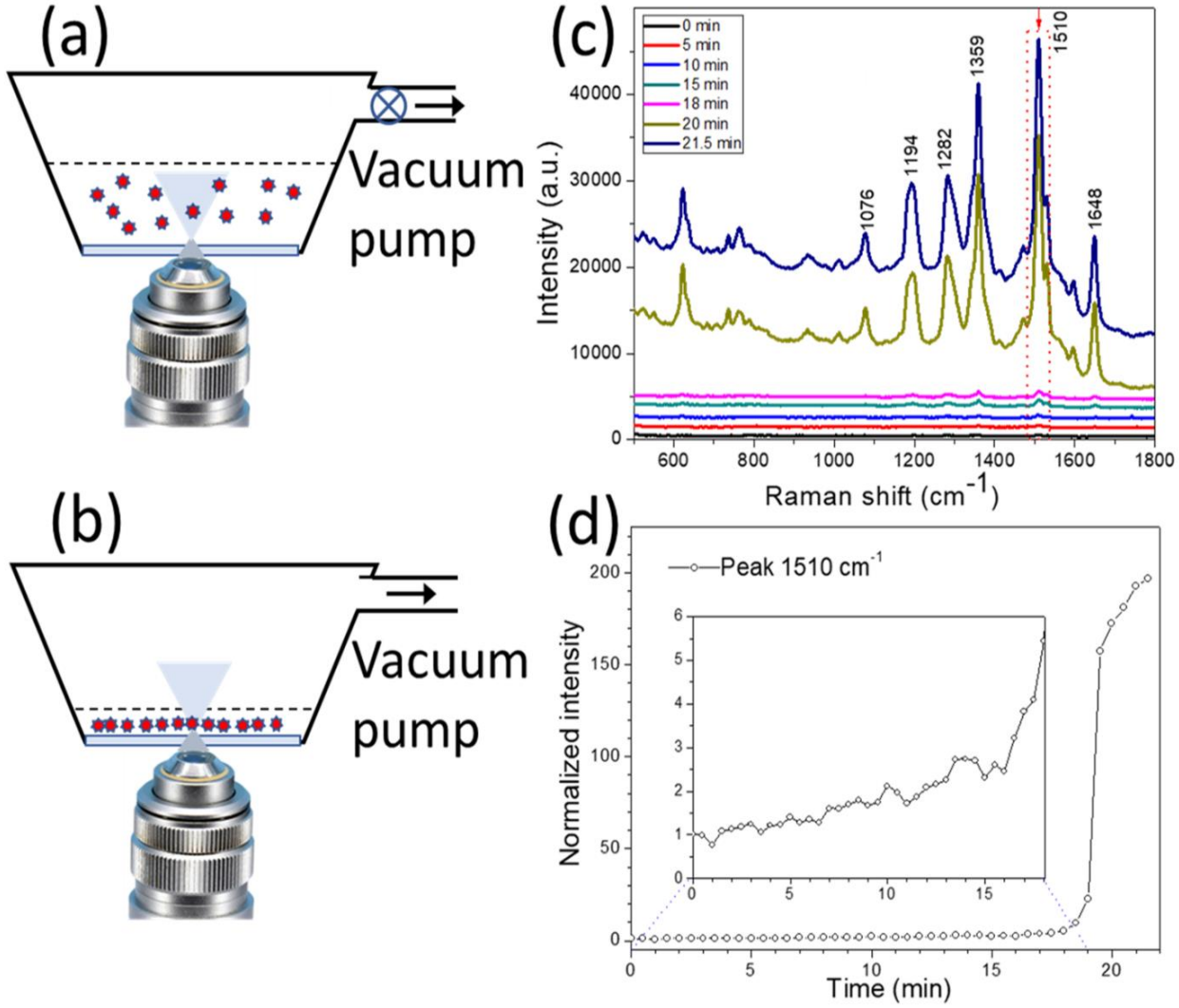

Figure 26. Demonstration of vacuum-enhanced micro-Raman spectroscopy of analytes in an aqueous solution. (a) Schematic of analyte molecules with a low concentration in the laser focus before vacuum evaporation; (b) Analyte molecules with a high concentration in the laser focus after vacuum evaporation. (c) Time-lapsed Raman spectra of 100 mL rhodamine B of 1 mM at various times after turning on the vacuum pump; (d) Normalized peak intensity of the 1510 cm-1 band vs time. The peak intensity was normalized to the peak intensity at t=0 min. The laser power was 20 mW at 780 nm and the acquisition time was 30s. The background spectrum of water was subtracted. An enhancement factor of ~200 was observed.

objective and confocally directed to a Raman spectrometer. About 100 μL of aqueous sample was placed to a sample holder with a bottom cover-glass diameter of 6.00 mm. Since the analyte molecules were dissolved in the entire solution, small number of analyte

molecules appeared in the laser focus region. After turning on the vacuum pump, the solvent liquid evaporates, and the number of analyte molecules inside the laser focal volume increases. After the complete evaporation of the solvent, there will be only a thin layer of analyte molecules on the bottom of the sample holder such that the enhanced Raman signal can be observed by focusing the laser beam on the bottom layer.

The enhancement of the VERS depends on the ratio of the concentrations of the analyte molecules before and after the vacuum evaporation of solvent solution. If C_v is the analyte concentration after the evaporation and C_o is the concentration before the evaporation, the enhancement factor will be C_v/C_o. For a given initial volume V, and the depth of laser focal volume of $2z_R$, the maximum enhancement factor (EF) is given by

$$EF = \frac{V}{\pi r^2 (2z_R)} \tag{30}$$

where $z_R = \pi w_0^2 n/\lambda$ is the Raleigh distance, w_0 the beam waist, λ the wavelength, r the bottom radius of the sample holder, and n the index of refraction. For the circular sample holder of 6.00 mm diameter, 100 μL of initial sample volume, and $\sim 1\ \mu m$ of the depth of focus, the maximum enhancement factor of ~3.5×10^3 was estimated.

The micro-Raman spectroscopy system consists of a laser excitation source at 780 nm (TEC-300–0780–0500, Sacher Lasertecknik), an 100X objective with NA of 1.3, an imaging spectrograph (Triax 320, Horiba Scientific). The sample holder had a bottom diameter of 6 mm covered with quartz glass and was connected to the vacuum pump for evaporation.

- **Samples Preparation**

The required materials, bovine serum albumin (BSA), rhodamine B, glucose, and ciprofloxacin, were purchased from Sigma Aldrich. The rhodamine B was dissolved in the distilled water to make a 1 mM stock concentration and further diluted for various applications. 500 mg/ml of stock concentration of BSA was first prepared and then diluted to 5, 4, 3, 2,1 and, 0.5 mg/ml for applications. 500 mM stock concentration of glucose was first prepared and then diluted to 10, 5, 2.5, 1, 0.5, 0.25 and, 0.1 mM with distilled water.

Ciprofloxacin is an antibiotic that is widely used to treat bacterial infections, and its residual may exist in human urine. To detect the amount of ciprofloxacin in urine, 2 mg/ml of ciprofloxacin was freshly prepared as the stock solution, and then, the ciprofloxacin was diluted with urine solution to 30, 20, 10, 5, and 2 μg/ml for Raman analysis.

The VERS was also combined with the SERS using the gold nanoparticles. For this experiment, 80 μl of 12.5μM rhodamine B was mixed with 20 μl of gold nanoparticle solution, thus forming 10 μM final rhodamine B concentration. Then, the 100 μl solution was taken in the sample holder for VERS analysis.

- **Synthesis of Gold Nanoparticles**

The gold nanoparticles were synthesized by the Lee and Meisel method[139]. The required chemicals, gold chloride ($HAuCl_4$:$3H_20$), trisodium citrate ($Na_3C_6H_5O_7$), and MilliQ water, were obtained from Sigma Aldrich. First, gold chloride (100 mL, 1 mM) was heated to boiling while stirring, followed by the addition of trisodium citrate (10 mL, 15 mM). After 15 s, the colour of the solution changed from yellow to deep red, and the solution was stirred for a further 50 min. The solution was then allowed to cool under

stirring, and the final AuNPs were stored at 4˚C until further use. The nanoparticles' resonance peak was measured using a UV-Vis spectrograph and found to be 535 nm.

- **Data Processing**

The data obtained from the spectrograph was transferred to the Origin Pro 2016 software, and the respective background was subtracted. Then, the spectrum was smoothed, and the baseline was corrected using SpectraGryph 1.2 (https://www.effemm2.de/). The final graph was plotted in Origin Pro 2016.

4.3 Results and Discussions

4.3.1 Detection of Glucose and Proteins at Low Concentration

The working principle of VERS is shown in Fig. 26(c,d) with rhodamine B. After turning on the vacuum pump, we recorded the time-lapsed spectra at 30s intervals for up to 21.5 min. At t=0 min (before turning on the pump), there was a tiny Raman signal. At the subsequent times after turning on the pump, the Raman signal of Rhodamine B molecules (e.g. the 1510 cm^{-1} band) increases due to solvent evaporation. At t=21.5 min, we observed a layer of the rhodamine B remained in the bottom quartz glass. Fig. 26(d) shows the normalized intensity of 1510 cm^{-1} during vacuum evaporation, which is normalized to the peak intensity at t=0 min. Before 18.5 min, the increase in Raman enhancement is slow as shown in the inset graph of Fig. 26(d). After 18.5 minutes, the Raman signal increased significantly, suggesting that more rhodamine B molecules were inside the laser focal volume. After 21.5 minutes, the largest Raman signal was observed as all water molecules

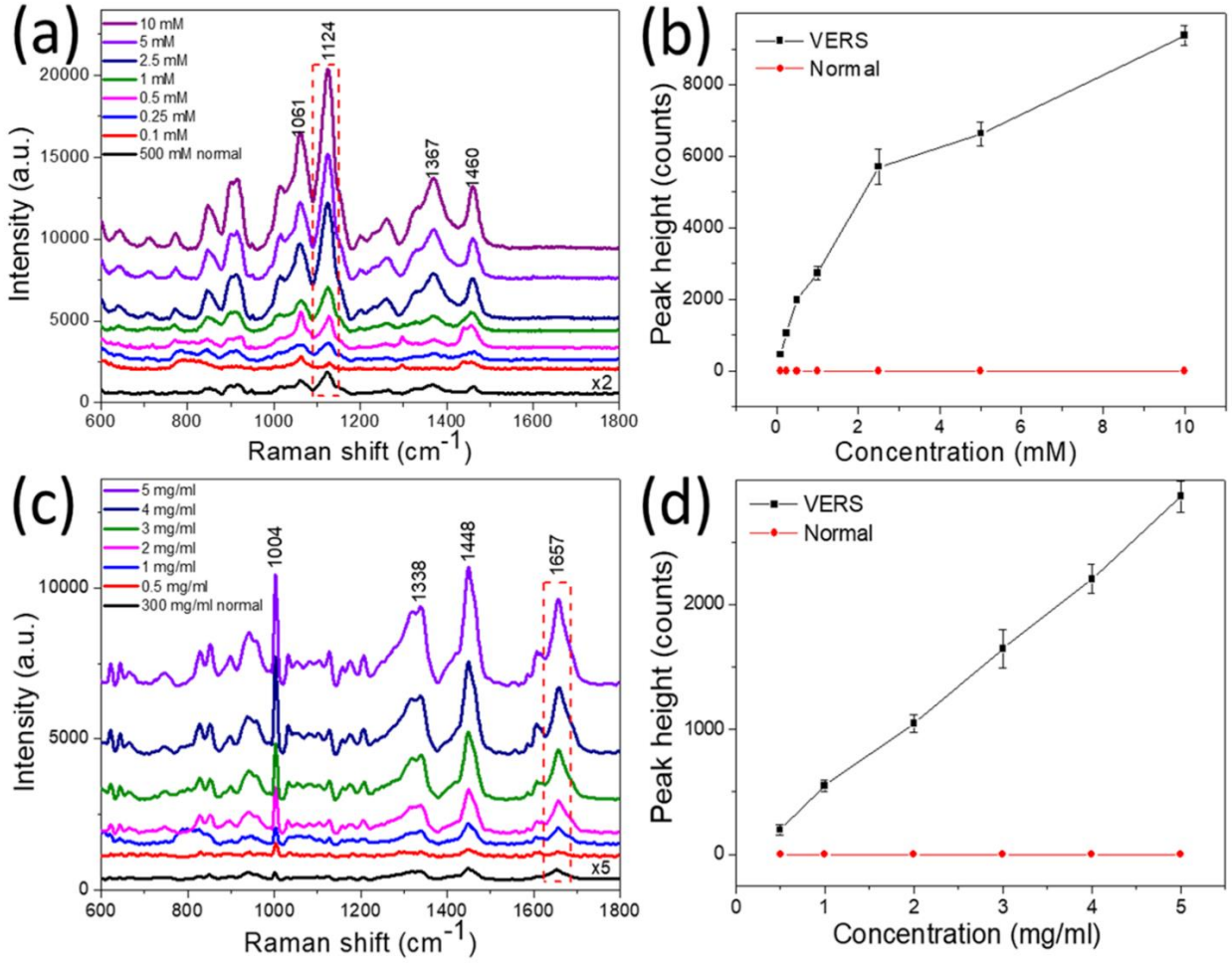

Figure 27. VERS and Normal spectra of glucose and BSA at various concentrations. (a) Normal Raman spectrum of 500 mM glucose and VERS spectra of 100 µL at 0.1, 0.25, 0.5, 1, 2.5, 5, and 10 mM glucose. (b) The peak intensity of the 1124 cm^{-1} band at different glucose concentrations. (c) normal Raman spectrum of 300 mg/ml BSA and VERS spectra of 50 µL at 0.5, 1, 2.5, 5, and 10 mg/ml BSA. (d) The peak intensity of the 1657 cm^{-1} band with different BSA concentrations. The laser power was 20 mW at 780 nm and the acquisition time was 30s.

were evaporated, and the concentration of rhodamine B molecules on the bottom quartz plate reached the maximum. ~ 200 times of enhancement was observed for this concentration compared to the initial concentration.

Fig. 27 shows VERS spectra of glucose and BSA at low concentrations after vacuum evaporation for 20 min. Fig. 27(a) shows the VERS spectra of the glucose of 0.1, 0.25, 0.5,

1, 2.5, 5, and 10-mM concentrations (with a volume of 100 μL) and the normal Raman spectrum of 500 mM glucose, respectively. The normal Raman signal was magnified two times, and all spectra were vertically shifted for clarity. The peak intensity of the 1124 cm^{-1} band at different glucose concentrations was measured and plotted in Fig. 27(b). By comparing the VERS spectrum of 0.5 mM glucose and the normal Raman spectrum of 500 mM glucose, we estimated the enhancement factor for glucose to be $2.0x10^3$, which is close to the theoretical prediction in Eq. (30). Also, we observed that the peak height of 1124 cm^{-1} was linear for low concentrations below 2.5 mM but was nonlinear for higher concentrations. Indeed, the VERY signal saturates for high glucose concentrations, because the thickness of the glucose layer after vacuum evaporation may larger than the depth of laser focal region so that only a portion of analyte molecules occupy the laser focal volume. Fig. 27(c) shows the VERS spectra of the BSA solution with of 0.5, 1, 2, 3, 4, and 5 mg/ml concentrations (with a volume of 50 μl) and the normal Raman spectrum of 300 mg/ml glucose, respectively. The peak intensity of the protein amide I band (1657 cm^{-1}) was plotted against the concentrations in mg/ml in Fig. 27(d). It was found that the VERS signal is linear with the studied concentrations Fig. 27(d), suggesting that we could apply the VERS technique for detecting low-concentrated proteins in aqueous solution.

4.3.2 Characterization of Ciprofloxacin Antibiotics in Human Urine

Fig. 28 shows the detection of ciprofloxacin in human urine using VERS. Ciprofloxacin is an antibiotic widely used for treating bacterial infections [120, 121]. After taking the antibiotic into the body, the remaining residual ciprofloxacin emerges through the urine. It is essential to know the amount of residual antibiotic remains in the body. Here, we used the VERS technique to detect the ciprofloxacin in the urine. First, the normal Raman spectrum of 2 mg/ml of ciprofloxacin dissolved in water was recorded and the background of distilled water was subtracted, showing the intense peak of ciprofloxacin at 1390 cm^{-1}. Second, the VERS spectrum of pure urine (with 0 µg/ml ciprofloxacin) was recorded, in which no ciprofloxacin peak (1390 cm^{-1}) was observed, and the strong peaks at 1003 cm^{-1} and 682 cm^{-1} are due to the urea compounds present in urine. This confirms that there is no any other analyte molecules contributing to the 1390 cm^{-1} band. Then, VERS spectra of ciprofloxacin concentrations of 0, 2, 5, 10, 20, and 30 µg/ml in urine were shown in Fig. 28(a), in which the 1390 cm^{-1} peak of ciprofloxacin was observed. A linear relationship between the ciprofloxacin 1390 cm^{-1} peak with studied concentrations was

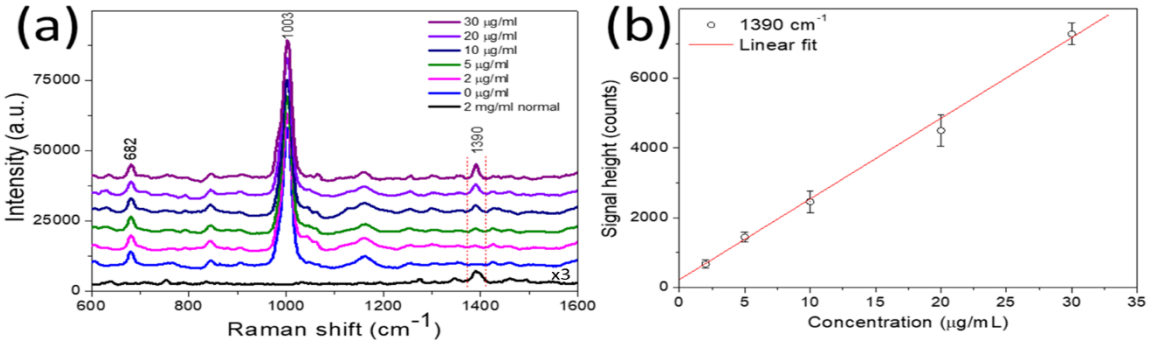

Figure 28. Detection of ciprofloxacin in urine using VERS. (a) Normal spectra of 2 mg/ml ciprofloxacin and VERS spectra of urine with ciprofloxacin concentrations of 0, 2, 5, 10, 20, and 30 µg/ml. (b) The peak height of 1390 cm^{-1} band verse ciprofloxacin concentrations. The laser power was 20 mW and the integration time was 120 s. When comparing the VERS spectrum of 30 µg/ml with the normal Raman spectrum of 2 mg/ml, an enhancement factor of ~ $4x10^2$ was achieved.

observed in Fig. 28(b), which may be used as a calibration curve for quantifying ciprofloxacin molecules in the urine in VERS. When comparing the VERS spectrum of 30 µg/ml ciprofloxacin with the normal Raman spectrum of 2 mg/ml, an enhancement factor of ~ $4x10^2$ was achieved. With VERS technique, ciprofloxacin at a low concentration of 2-5 µg/ml can be detected in human urine, although strong urine background at 1003 cm^{-1} was present.

4.3.3 **Detection of Biomolecules Resolved in Solvents with Intense Raman Background**

Many chemical and biological molecules cannot be easily dissolved in water, but they are easily dissolved in organic solvents like Dimethyl sulfoxide (DMSO). However, DMSO solvent has very strong Raman background that suppresses the weak Raman signals of analyte molecules, especially at low concentrations. Here, we show that the VERS technique allows the reduction in the solvent background due to vacuum evaporation and thus allows the detection of analyte molecules.

ARS-1620 is an inhibitor molecule for Ras proteins, and DMSO is usually used as the solvent to dissolve this molecule. Fig. 29(a) shows that the solvent DMSO has very strong Raman background (curve A) that significantly suppress the observation of ARS-1620's Raman signal (curve B). However, by the evaporation of DMSO solvent, we can easily detect the Raman spectrum of ARS-1620 (curve C), suggesting that the VERS technique can effectively reduce the solvent Raman background. Similarly, we cannot observe Raman spectrum of 0.1 mM rhodamine B dissolved in DMSO, as shown in curve B in Fig. 29(b).

However, we can easily detect the rhodamine B signal (curve C in Fig. 29(b)) by the VERS technique to remove the high Raman background of DMSO solvent.

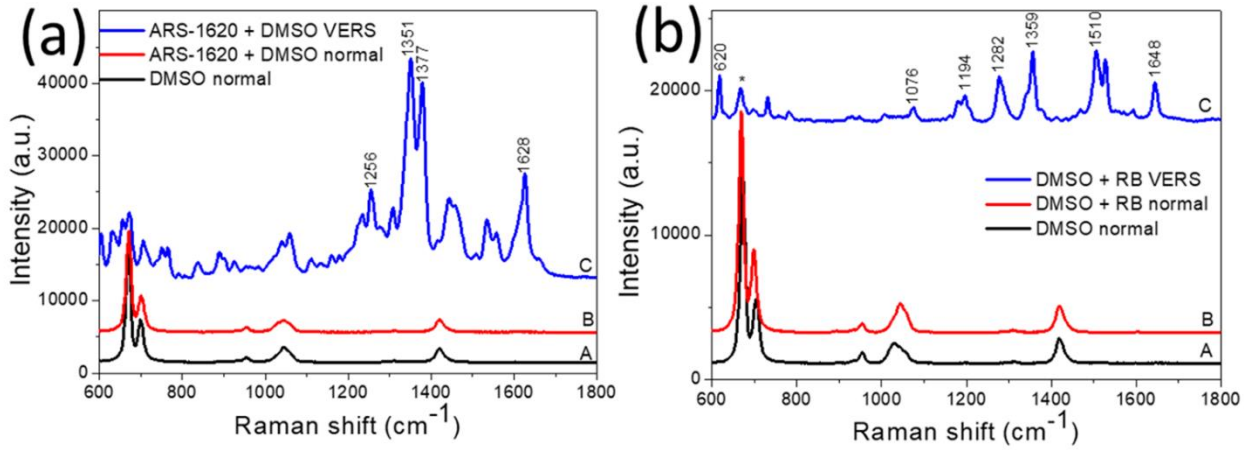

Figure 29. (a) Detection of ARS-1620 molecules dissolved in DMSO solution using vacuum-enhanced micro-Raman spectroscopy. Curve A is the Raman spectrum of pure DMSO and Curve B is the normal Raman spectrum of 10mM ARS-1620 dissolved in DMSO with 1s integration time. Curve C is the VERS spectrum of 10mM ARS-1620 dissolved in DMSO with 60s integration time. (b) Detection of 100 mM rhodamine B dissolved in DMSO solution. Curve A and B are the normal Raman spectra of the pure DMSO and rhodamine B in DMSO, respectively, with 1s integration time. Curve C is the VERS spectrum of rhodamine B dissolved in DMSO solution with 30s integration time. The laser power was 20mW. The symbol '*' denotes the remaining contribution from DMSO.

4.3.4 Dependence of the Enhancement Factor

As shown in Eq. (30), the enhancement factor (EF) of the VERS technique can be increased by reducing the radius r of the sample holder and increasing the initial volume V of liquid sample. Higher EF value will increase the detection limit. Fig. 30(a) shows the VERS spectra of 0.5 mM glucose with the sample holder of 3.0 mm and 6.0 mm in diameter, respectively, while keeping the initial volume of sample used and other parameters the same. While the EF was ~$2x10^3$ for the 6.0-mm sample holder, the EF for the 3.0-mm sample holder was ~ $8x10^3$ for 0.5 mM glucose because the signal level for the 3 mm sample holder was about four times higher from Fig. 30(a). This finding indicates

that decreasing the diameter of the holder yields higher VERS enhancement for a given analyte at low concentration, since a small sample holder decreases the base area and thus increases the thickness of the analyte layer so that there will be more analyte molecules in the laser focal volume.

Fig. 30(b) shows the dependence of VERS enhancement factor on the initial volume of liquid. For this experiment, 25, 50, 100, and 150 μL of 0.5 mM glucose were taken, respectively, and VERS was performed. The peak height of the 1124 cm^{-1} band of glucose molecules was measured and plotted as the function of the sample volume in μL. The observed VERS signal height is linearly proportional to the volume of the liquid taken, as predicted in Eq. (30).

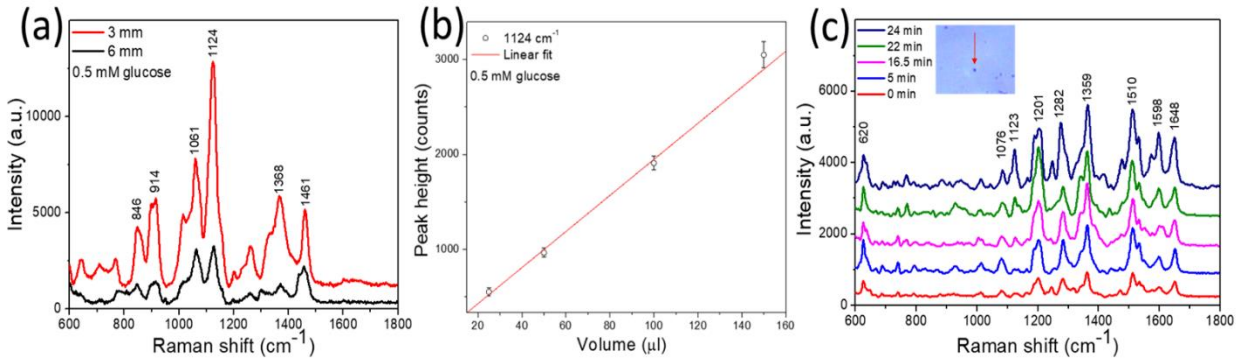

Figure 30. (a) VERS spectra of 0.5 mM glucose with a sample holder of 3.0 mm or 6.0 mm in diameter. (b) The dependence of the peak intensity of glucose's 1124 cm^{-1} band on the initial volume of the liquid sample taken. The laser power was 20 mW with the integration time of 30s. (c) Combination of the VERS and SERS of rhodamine B with AuNPs. The time-lapsed Raman spectra of a single Au nanoparticle immersed in 10 μM rhodamine B solution. The laser power was 35 mW with 1s integration time.

The VERS may also be combined with the SERS method to further increase the enhancement factor and thus increase the detection limit. For this experiment, we mixed 10 μM rhodamine B with gold nanoparticles. After loading the mixed solution onto the sample holder, we observed the gold nanoparticles sitting on the quartz plate under the microscope.

To quantify the enhancement factor, we focused on a single nanoparticle fixed on the bottom quartz plate, and the SERS spectra from the same nanoparticle was recorded before and after turning on the vacuum evaporation at different times. As shown in Fig. 30(c), it can be observed that the VERS signal from a single AuNP (e.g. the spectrum at 24 min) is increased upon the vacuum evaporation, comparing to the SERS signal (the spectrum at 0 min). The enhancement factor for SERS only was estimated to be $3x10^5$. The VERS from the same nanoparticle enhanced the SERS signal by a factor of 4. Hence, the total enhancement factor with the combination of VERS and SERS was ~ $1.2x10^6$.

We have measured BSA proteins, glucose, and ciprofloxacin in solution using the VERS technique. It should be noted that the concentration used in these experiments was not the detection limit for these analyte molecules. Although we showed the detection of ciprofloxacin up to 2 μg/ml using the VERS technique, we may increase the sensitivity using longer integration time, smaller bottom size of the sample holder, and larger initial volume of the sample solution.

4.4 Conclusions

We have developed a simple and easy-to-use vacuum-enhanced micro-Raman spectroscopy (VERS) for the detection of chemicals, biomolecules, and antibiotics molecules at relatively low concentrations in aqueous solution. The VERS technique does not require any sensor materials to enhance the Raman signal but relies on the increase in local molecular concentration within the laser focus using vacuum evaporation of solvents and does not alter the Raman spectra. We demonstrated that an enhancement factor of ~10^3-10^4 was observed for glucose and protein samples and the enhancement factor depends on

the size of the sample holder and the volume of the liquid sample used. The VERS can be used to detect ciprofloxacin antibiotics in human urine at a level of 2 μg/ml. The VERS technique is particularly useful for detecting biomolecules resolved in a solvent such as dimethyl sulfoxide (DMSO) that has an intense Raman background due to the evaporation of the solvent. when combined with SERS, this VERS can further increase the detection limit of the analyte molecules in solutions. Although the motivation to develop this VERS technique was to apply for ras proteins isoforms, we were unable to apply this technique due to the limited Ras protein sample availability.

Chapter 5. SERS and SERS-based Super Resolution Raman Imaging

5.1 Introduction

To detect the sample at single molecule level, surface-enhanced Raman scattering (SERS) technique would be useful because of large enhancement factor. At single molecule level with SERS, it was reported in the literature that the signal intensity fluctuation occurs. The origin of fluctuations in single-molecule SERS remains unclear. Various factors contribute to SERS enhancement at the nanometer scale, including the clustering of nanoparticles, the positioning of analyte molecules on the nanoparticle surface, laser power, accumulation time, and environmental conditions such as temperature and pressure [52]. Typically, when the Raman signal is accumulated over an extended period, the fluctuation is averaged out, making them difficult to observe. However, with a short accumulation time, the intensity fluctuations become significant. In this study, we examine the effects of variations in laser power and pressure on Raman intensity fluctuations using SERS-based Raman imaging. This will help to study the behavior and dynamics of analyte molecule like Ras protein when binding to nano particle at single molecule levels. In the present work, we tested the feasibility of this technique in our laboratory using R6G molecules and silver nanoparticles. Silver nanoparticles produce a significantly higher EM field enhancement at the NP surface than gold nanoparticles in visible excitation wavelength. This will enhance the Raman signal more and beneficial for the single molecule detection [122].

5.2 Materials and Methods

This section describes the samples used, the method of preparation, and the data analysis technique in this experiment.

5.2.1 Experimental Setup: Raman Imaging Microscope and Micro-Raman Spectroscopy System with 532-nm Excitation Laser

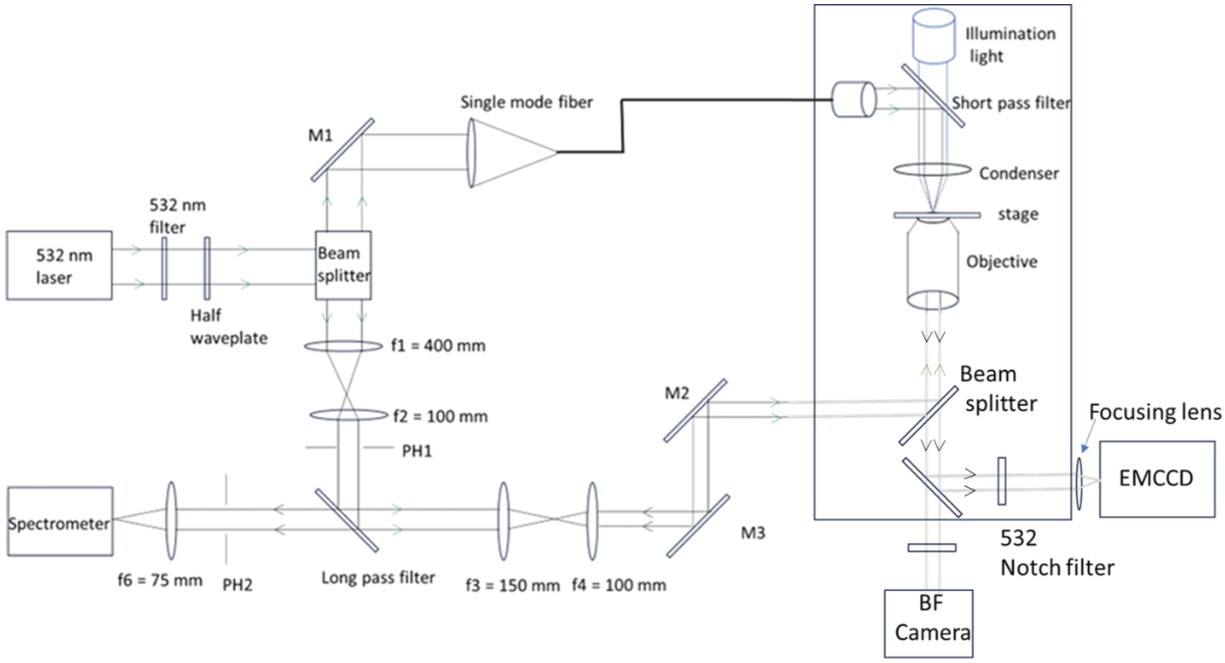

Figure 31. Schematic diagram of Raman imaging/spectroscopy system using 532-nm laser.

The schematic system for the Raman imaging/spectroscopy system is shown in Fig. 31.

The major components are described below.

- **Excitation laser.** A diode-pumped solid state (DPSS) 532-nm green laser (Thorlabs Inc, #DJ532-40) is used as the excitation source. The laser beam first passes through a narrow bandpass filter at 532 nm (Semrock Inc. #LL01-532-12.5), which removes other wavelengths of light and only allows 532 nm light to pass through. A half-wavelength waveplate changes the polarization of the incident laser beam, enabling the adjustment of laser intensity when the beam is divided with a polarization cube beam-splitter (PBS). The beam splitter divides the laser beam into two beams; one is used for the excitation for Raman imaging from the top of the microscope via a

single-mode optical fiber, and the other serves as an excitation source from the bottom of the microscope.

- **Microscope**. An Olympus IX81 microscope was used with a 100X and 1.49 NA objective, in which the illumination light was replaced with a high-power blue light-emitting diode (LED) at 470 nm and 5W. For Raman imaging excitation, the laser is fed to the single-mode fiber using a lens with 4.5 mm focal length. The output of the fiber is mounted to the top of the microscope, and the 532nm laser is directed through a short-pass dichroic mirror (Semrock, #FF505-SDi01-25x36, R_{avg} > 98% for 513 – 725 nm; T_{avg} > 90% for 446 – 500 nm), which transmits the blue LED light but reflects the 532 nm green laser for Raman imaging.
- **Bright-field imaging camera**. To view the bright-field image of the hot nanoparticles (80nm silver NPs) illuminated by the blue LED light at 470nm, we used a 16-bit digital CCD camera (QSI 520 CCD camera) with a bandpass filter (semrock, #FF01-466/40-25), which has a transmission window from 446 nm to 486 nm. This filter only allows the bright field images to pass and blocks the excitation laser.
- **Raman imaging camera**. The diameter of the Raman imaging laser is approximately 20 micrometers under the microscope, which excites the SERS from a few Ag NPs within the field of excitation (~20 μm). The scattered Raman signals are collected with the same objective (100X and 1.49 NA). The Raman signal is divided by a broadband beam splitter (DM1) with 60/40%, where 60% of the signal is sent to an EMCCD camera (Andor, iXon Ultra 897) to record the Raman image, and 40% is sent to the spectrometer to record the Raman spectrum. A 532 nm notch filter is placed in the front of the EMCCD to block the excitation laser at 532nm, ensuring that only the Raman signal with the wavelength greater than 550nm is recorded for imaging. To further reduce effect of the fluorescence signal, we added a bandpass filter (Semrock, #FF01-560/25-25)

with a bandwidth of 545-575 nm in the front of the EMCCD camera, which allows the detection of Raman signals from 448 cm^{-1} to 1405 cm^{-1} for Raman imaging when excited with a 532nm laser.

- **Raman Spectroscopy**. The other part of the laser serves as the bottom excitation source. After the beam splitter, this portion of the laser is narrowly focused using a pair of convex lenses with focal lengths of 400 mm and 100 mm, respectively. The laser is then reflected by a long-pass filter (Semrock, #LPD01-532RS-25) at 45 degrees, which reflects the 532nm laser but transmits the Raman scattering light. The collimated laser beam subsequently passes through another pair of lenses with focal lengths of 150 mm and 100 mm, which are used to adjust the laser focus exactly on the imaging plane of the microscope. After the reflections from mirrors M2 and M3, the laser strikes a dichroic mirror (DM1), enters the microscope objectives, and focuses on a single hot nano-particle on the sample slide. The Raman scattering light from the excited nano-particle is collected with the same objective and directed back to the long-pass dichroic filter and focused into the entrance slit of an imaging spectrometer (HORIBA, TRIAX 320). A cryogenic CCD detector (HORIBA, Symphony 1024 x 256 open electrode) is used for the acquisition of Raman spectrum. An ultra-steep long-pass edge filter (Semrock, #LP03-532RS-25) is placed in the front of the entrance slit of the spectrometer to reject the elastic scattering light at 532 nm but transmit the Raman scattering light (with OD>6 for 532nm and T_{avg}>93% for 535.4-1200nm).
- **Resonance Raman Spectroscopy.** The intensity of Raman scattering strongly depends on the wavelength of the incident light. Generally, shorter excitation wavelengths lead to stronger Raman scattering signals because the intensity of Raman scattering is proportional to $(1/\lambda)^4$, where λ is the wavelength of the excitation light [91]. Some molecules exhibit resonance enhancement (such as Rhodamine 6G molecules at 532 nm), where their Raman scattering intensity is significantly

enhanced when the excitation wavelength matches an electronic transition of the molecule. In such cases, using a wavelength that is resonance with the molecule's electronic transitions can dramatically increase the Raman scattering signal. This is very useful when measuring a low concentration sample or at single molecule levels. In this experiment, we chose Rhodamine 6G as the analyte molecules to test the SERS with Ag NPs and Raman imaging, which has a resonant absorption at 532 nm and thus shows resonance enhancement in Raman scattering in addition to surface enhancement by Ag NPs.

- **Micro-Vacuum Chamber.** A miniature vacuum chamber is home-made to mount on the stage of the Olympus IX81 microscope. A glass coverslip (with a thickness of 80-100 μm) was sealed as the bottom plate of the vacuum chamber. The Ag NPs and Rhodamine 6G are deposited on the inner surface of the coverslip of the vacuum chamber. The pressure of the vacuum chamber is pumped with a rotary vane pump (Agilent #DS302) and measured with a thermocouple gauge tube (LDS Vacuum, #200TC-TORR-531). The pressure of the micro-vacuum chamber can be changed from 1 atm (760 Torr) to 4 mTorr with a vacuum valve (LDS Vacuum, #MRV-NW25).

5.2.2 Sample Preparation

The chemicals used in this experiment included 80 nm silver nanoparticles, rhodamine 6G, and sodium chloride (NaCl). The 80 nm silver colloidal nanoparticle solution was purchased from Sigma Aldrich and used as received, without further processing. Rhodamine 6G powder was also obtained from Sigma Aldrich. A stock concentration of 1 mM rhodamine 6G was freshly prepared in distilled water, which was then diluted to 10 nM in deionized (DI) water. Additionally, a stock solution of 1 M NaCl was freshly prepared and subsequently diluted to 33 mM for the experiment.

To perform the measurement, a sample with a nano-mole level concentration was prepared to approximate single-molecule level concentration. We mixed 2 µl of a 10 nM R6G solution with 4 µl of silver nanoparticles and vortexed the solution for 30 seconds. This vortexing aids in attaching the R6G molecules to the silver nanoparticle surfaces. Subsequently, 4 µl of a 33 mM NaCl solution was added to the mixture, and it was vortexed for another 30 seconds. The addition of NaCl promotes the clustering of nanoparticles, creating hotspots.

Next, a small drop of the liquid mixture was placed on a quartz glass sample holder to form a thin film of R6G and silver nanoparticles. We used quartz glass and low-fluorescence oil to minimize the fluorescence contributions to the background signal. To prepare the thin film, 1 µl of the prepared solution was deposited on the quartz sample holder and dried under vacuum until a thin film formed.

The sample was then washed with fresh distilled water to remove the NaCl and any unattached R6G by gently flowing water over it. After drying the sample again to remove any remaining water, it was placed on the microscope objective for measurements. The imaging laser illuminated the sample from above and scanned the thin film on the glass plate. Through this process, we could observe Raman imaging of the nanoparticles containing hotspots. Not all hotspots showed fluctuations, but we specifically selected those that did for data collection. This approach allowed us to identify the hotspots and collect the necessary data.

5.2.3 Data Collection

The sample was placed under the microscope and illuminated with a top imaging laser. We located the fluctuating hotspot using the EMCCD camera and then excited that nanoparticle using bottom excitation lasers. During this process, the top laser was turned off. After the bottom laser

excitation, we recorded the Raman imaging from the EMCCD camera, the Raman spectrum from the spectrometer, and the bright field image.

5.2.4 Data Analysis

The recorded images were first processed in ImageJ, and then super-resolution images were created using ThunderSTORM software. The Raman spectrum was originally processed, and the scatter plot diagram with intensity variation was plotted in MATLAB.

- **Drift Correction**

We recorded the bright-field images for 40 seconds at a rate of ~6 frames per second. The bright-field image of a single nanoparticle appears dark at the center with a diffraction-limit spot size (see Fig. 32(A)). To prevent sample drift, we locked the vertical direction using our microscope's ZDC-stage lock to stabilize the stage drift in z direction. We also used a clamp to reduce the sample stage's motion in the x and y directions. In addition, we utilized the drift correction features in ThunderSTORM software to eliminate the remaining drift motion in x-y directions. As a reference for this procedure, we selected a non-hot nanoparticle close to the studied hot particle. ThunderSTORM then tracked this nanoparticle for every image frame and subtracted this movement from the data, thus eliminating the drifts. Fig. 32 shows the steps to correct the drift motion using the bright field images. The nanoparticle looks dark in the bright field image, and we inverted this image so that the center of the nanoparticle is bright. Then, we apply drift correction features in ThunderSTORM software, in which we have the option to track which particle in the software. We selected the non-hot nanoparticle for tracking the image drifts and can save the data as a correction factor. Then, we track the actual hot particles and apply the correction factor. In this way, we can remove the drift over time.

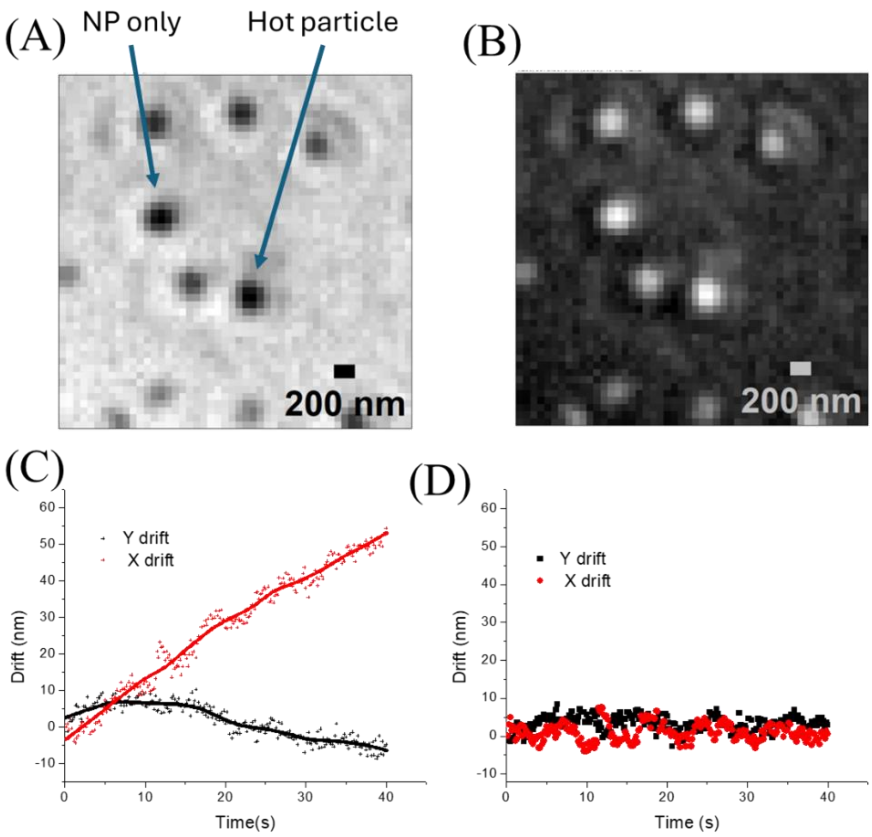

Figure 32. Correction for stage drifts in ThunderSTORM software. (A) The acquired bright field image, (B) the inverted image of the bright-field image, and (C) the drift of the stage during 40s interval time, (D) the image drift of the hot particle after the drift correction.

5.3 Results and Discussions

5.3.1 SERS Raman Fluctuations are Observed at Single Molecule Levels

The typical Raman spectrum of rhodamine 6G and the Raman image of the hot spot containing silver nanoparticles and R6G is shown in Fig. 33. The Raman image in Fig. 33(B) only includes the Raman spectral peaks between the two red dotted lines in the spectrum shown in Fig. 33(A). A bandpass filter was used in the front of the EMCCD camera to remove the background outside 560 ± 15 nm, allowing only signals from the 545 nm to 575 nm range to pass through to form the Raman image. The Raman spectrum was obtained using 1μW of the bottom laser with

an accumulation time of 1s. The Raman image was recorded using a top laser for 200 ms, with a laser intensity of 1 kW/cm^2.

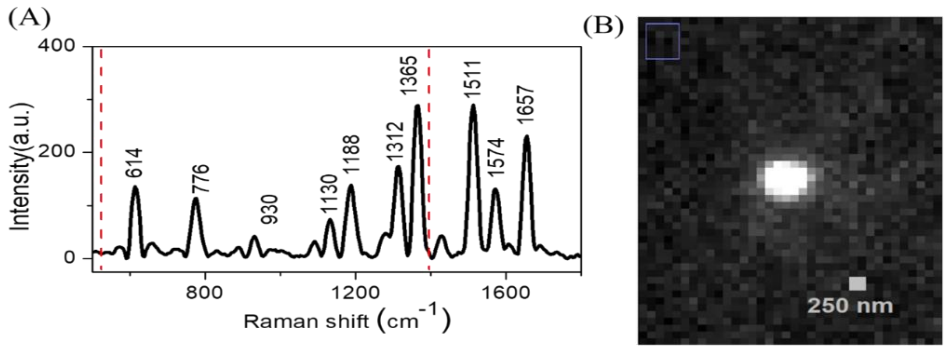

Figure 33. (A) Raman spectrum of R6G acquired with the bottom excitation laser of 1 µW power and with the spectrometer CCD with an accumulation time of 1 second. (B) Raman image of the R6G using silver nanoparticle acquired for 200 ms with the EMCCD camera.

One important property of SERS of the hot particle with analyte at a single-molecule level is the fluctuation of the Raman intensity as the function time, even though the hot particle is excited with a constant laser power. To identify if the intensity fluctuations in Raman imaging and in Raman spectrum are the same time-course, we first identified a hot particle using the EMCCD camera when excited by the top laser, which showed an extremely bright in the Raman image. Once the hot particle was identified, the top laser was turned off and then the hot particle was excited using the bottom laser with ~2 µW power. The Raman spectra were recorded every 1s for 40 seconds, and simultaneously, Raman imaging and bright field images were recorded every 200 ms for 40 seconds.

Fig. 34(A) shows the intensity fluctuations in the Raman spectra and Fig. 34(B) shows Raman images at different times. We have plotted only a few spectra to show the SERS behavior of the hot particle at single-molecule levels. As seen in Fig. 34(A), at t=1s, we observed an intense R6G's SERS signal, and at t=10s, this Raman signal completely disappeared. We observed the

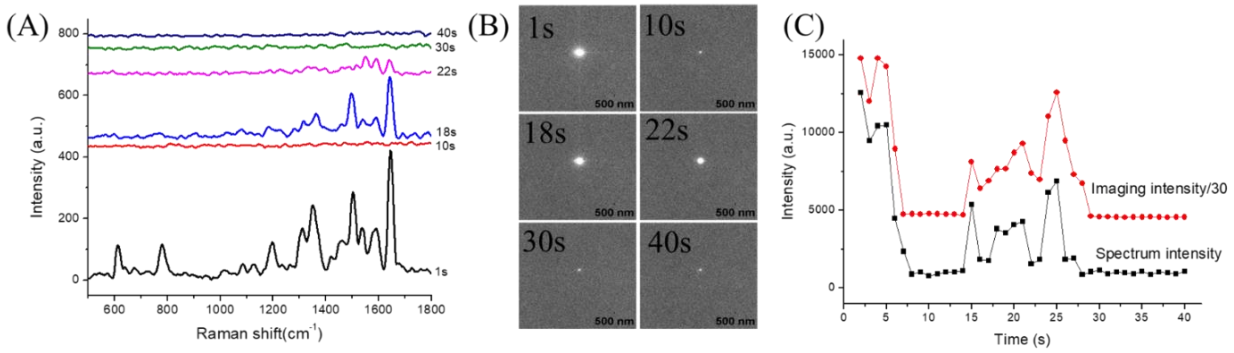

Figure 34. (A) Time-lapse Raman spectra show the Raman intensity fluctuations. The Raman spectra of a hot particle were acquired at different times using 2.2 μW laser power and 1s accumulation time. (B) Raman images of the hot particles at different times. (C) Raman image intensity and Raman spectrum intensity vs time. The red line is the summed intensity from the Raman image recorded using an EMCCD camera. The black line is the sum of the Raman spectrum's intensity of all peaks from 500 to 1400 cm^{-1} recorded using the spectrometer's CCD.

Raman signal back from t=15 to 30 s but with time-varying intensity at different times. For example, the spectrum at t=18s shows the re-appearance of the Raman signal but with lower intensity, and the Raman signal at t=22s is lower. Fig. 34(C) shows the comparison of Raman spectrum's intensity with Raman image's intensity. From this figure, we can observe that these intensity fluctuations are highly correlated. The Raman image's intensity was divided by a factor 30 for visual clarity in the graphs.

5.3.2 SERS Intensity Fluctuations at the Atmospheric Pressure and 4 mTorr Pressure

Since the mechanism of the SERS intensity fluctuation remains unclear, we intend to test if the SERS intensity fluctuation is dependent on the environmental pressure where the hot particles locate. The SERS Raman fluctuation behaviors of two different hot particles was studied at the atmospheric pressure (760 Torr) and at a low pressure of 4 mTorr using a top laser with an intensity of 1 μW/cm^2. First, we recorded the time-lapse Raman images of the hot particles at the atmospheric pressure at a rate of 17 frame per seconds for 45 seconds so that their SERS intensity

fluctuations can be measured. Second, the pressure of the vacuum chamber was reduced to 4 mTorr by turning on the vacuum pump, and the Raman images of the same hot particle were recorded for

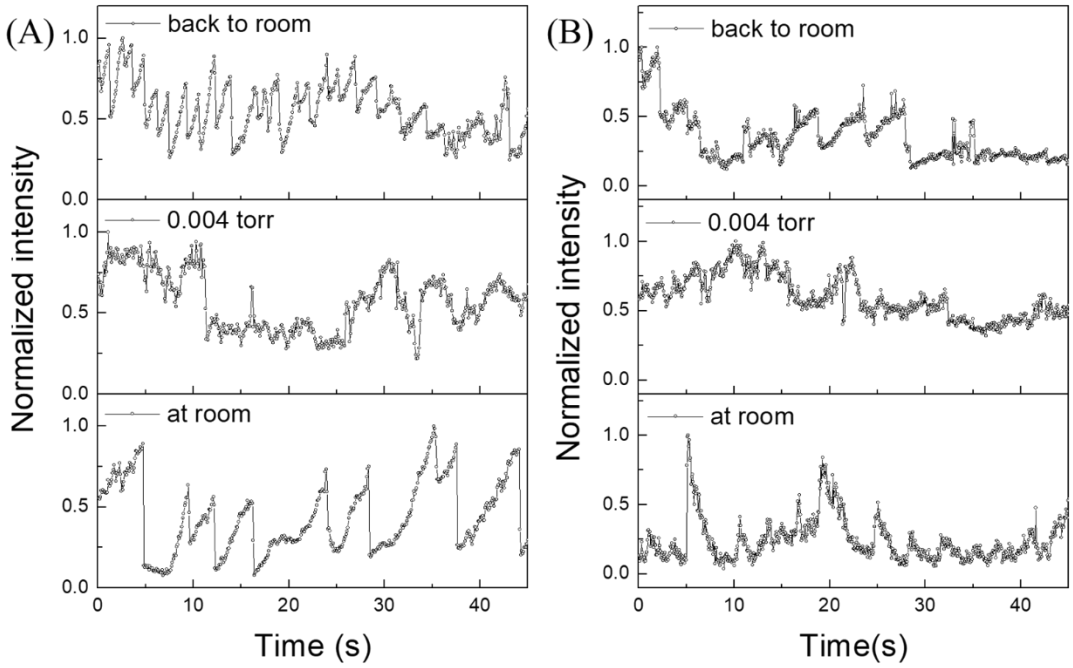

Figure 35. The change of Raman intensity of hot particle vs time at the room pressure (760 Torr) and at a low pressure of 4 mTorr for two hot particles (A) and (B).

another 45 seconds. After this, the vacuum pump was turned off, the pressure was returned to the atmospheric pressure, and the Raman images were recorded again for 45 seconds. Once all the images were captured, we applied ThunderSTORM and obtained the intensity of each Raman image. This intensity was then plotted against time, as shown in Fig. 35. From both graphs, it is evident that the SERS intensity fluctuations are more pronounced at the atmospheric pressure. When the pressure is reduced to 4 mTorr, the SERS intensity fluctuations decrease.

5.3.3 SERS Intensity Fluctuations at Different Laser Powers

The SERS intensity fluctuation behavior was studied at various laser excitation powers from the bottom laser. All the Raman images were acquired using an EMCCD camera. Initially, we used a 0.9 μW laser power and recorded the Raman images of a hot particle for 45 seconds at a rate of 5 frame per second and the exposure time per image frame was 200 ms. We then increased the laser power to 2.2 μW and recorded the Raman images for another 45 seconds. This procedure was repeated for 4.4 μW and 8.9 μW, respectively. We used ThunderSTORM to obtain the

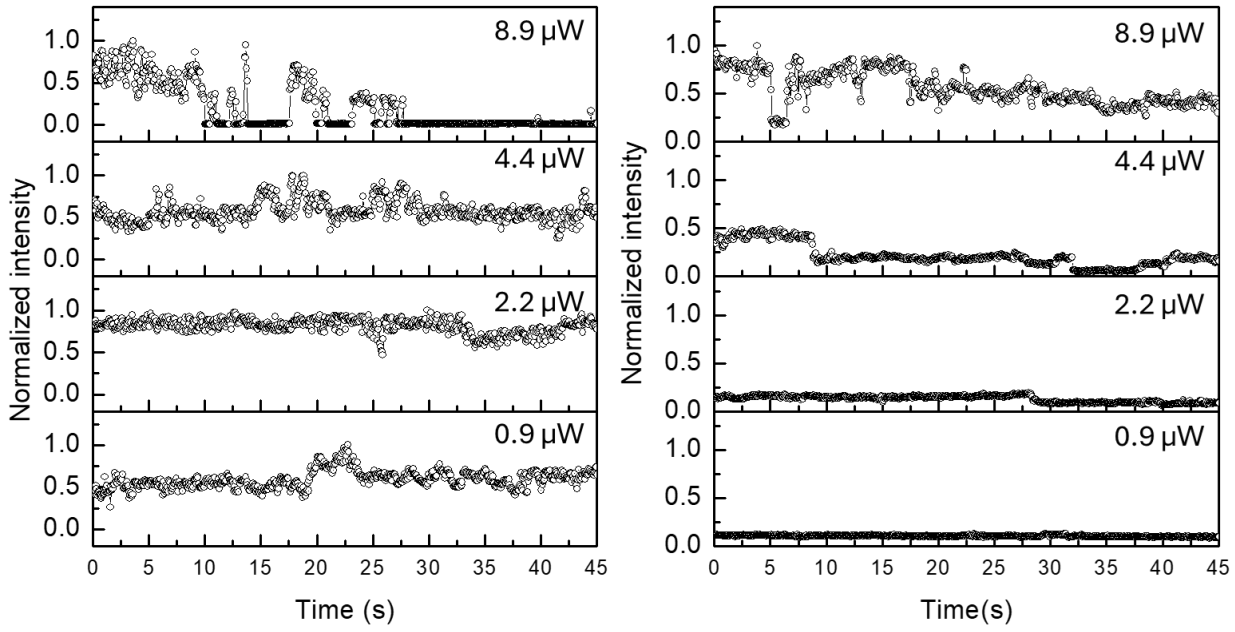

Figure 36. Fluctuations of Raman intensity at different laser powers for two different hot particles (A) and (B).

intensity of each Raman image and plotted it as the function of time. Fig. 36 shows that at lower laser powers, the intensity fluctuations are minimal, and there is less “blinking”, which means the intermittent appearance and disappearance of the SERS signal. At a higher laser power of 8.9 μW, the fluctuations are greater, and there is more pronounced blinking. During the blinking period, the nanoparticle is not “hot” probably due to geometry change of the cluster of nanoparticles or

the analyte molecules (R6G) moving off the “hot” position of the nanoparticles, both leading to the enhancement in Raman scattering signal not sufficient large to observe. We also attempted to use a 16 μW laser power, but the signal quickly disappeared and did not return for a couple of minutes. The higher laser power might destroy the geometry of the hot particles due to the high temperature.

5.3.4 Hot Spot Tracking Using Bright Field Image

We recorded the bright field images of hot particles to observe the SERS emission and the change of hotspot locations. We selected a hotspot and used a stable cold nanoparticle or nanoparticle cluster as the reference particle to correct the stage drift. Fig. 37 illustrates the methods used to generate the plots of center position of the hot particle, utilizing bright field images to track the location of the nanoparticles' emission centers.

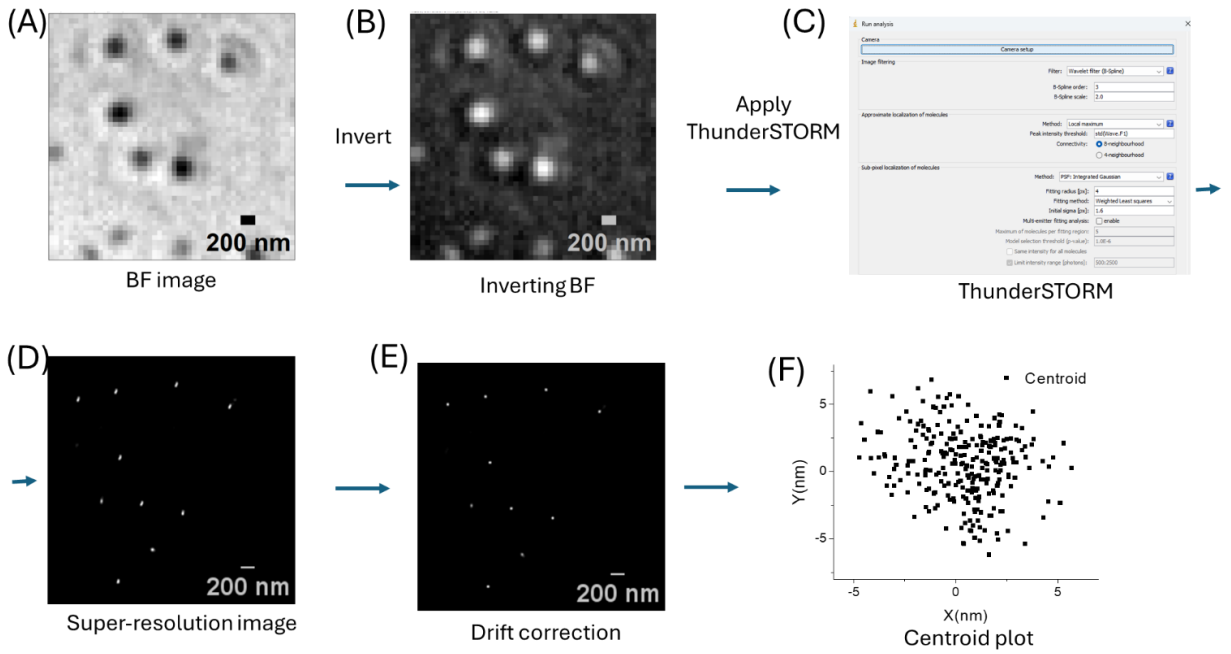

Figure 37. Methods to find the centroid of bright field image using ThunderSTORM software. (A) Bright field image, (B) the invert of bright field image, (C) ThunderSTORM software, (D) super resolution image of bright field image (A), (E) the centroid image after drift correction, and (F) scattered plot of the center coordinate of one hot particle for each frame.

Fig. 37(A) shows the bright field image captured by the imaging camera, with each image having an exposure time of 200 ms and acquired over a 45-second period. The pixel size of the bright field camera was ~73.5 nm and the bright field image was diffraction-limited with a spot size ~300 nm. Fig. 37(B) shows the inverted image of Fig. 37(A), highlighting the center of the nanoparticle. We then applied ThunderSTORM software as depicted in Fig. 37(C) to obtain the super-resolution image of the center coordinates of each particle shown in Fig. 37(D). During the 45-second recording, stage shifts might occur due to thermal drifting or environmental vibrations. ThunderSTORM includes features that correct drift motion based on the first frame. ThunderSTORM software can also provide the intensity and the x and y coordinates of each nanoparticle's center. We applied this drift correction to the super-resolution image to compensate for such shifts, as shown in Fig. 37(E). By plotting these coordinates, we generated the scatter plot of the emission centers of one particle, as shown in Fig. 37(F).

Fig. 38 displays the centroid plot of the emission center of two hot particles obtained from the bright field images when excited at different laser powers. Once the hot particle was identified

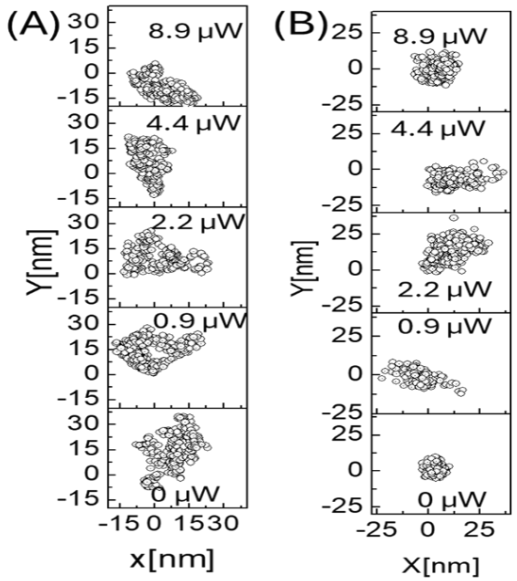

Figure 38. Centroid plot of two hot nanoparticles extracted from the bright field images.

using Raman imaging, the laser was turned off, and the control bright field image was acquired using bottom laser at 0 μW of laser power. The same hot particle was then excited using the bottom laser powers of 0.9, 2.2, 4.4, and 8.9 μW and bright field images were acquired at each laser power. The images were then analyzed as discussed in Fig. 37 and the centroid of each frame was plotted, as presented in Fig. 38. The data in Fig. 38(A) and (B) show that the location of the cluster with the application of excitation laser slightly deviates from the location without the excitation power (0 μW). This deviation may be due to minor changes in the configuration of the hot nanoparticle clusters following the application of the laser, probably due to the high temperature generated on the hot particle.

5.4 Conclusions

This comprehensive study explored the dynamic behavior of surface-enhanced Raman scattering (SERS) using hot particles, focusing on the influence of environmental conditions and laser power on the stability and intensity fluctuation of SERS Raman signals. By examining the fluctuations in Raman intensity at atmospheric pressure (760 Torr) and low pressure (4 mTorr), it was observed that fluctuations were more pronounced at atmospheric pressure, highlighting the impact of atmospheric conditions on single molecule SERS signal stability. The experimental data further revealed that lower laser powers resulted in smaller fluctuations and less blinking, suggesting a more stable interaction between the laser and the hotspots. As the laser power is increased, particularly at 8.9 μW and above, there is a notable increase in intensity fluctuations and blinking, potentially due to thermal or photo-induced changes in nanoparticle clusters. The use of bright field and Raman imaging for tracking hotspots together with ThunderSTORM software, was effective in accurately identifying and analyzing the change in hot particle's center

position. Collectively, these findings underscore the necessity of precise experimental control over both laser parameters and environmental conditions to optimize SERS applications at single molecule levels. This study not only advances our understanding of SERS behavior at the single-molecule level under varying experimental conditions but also emphasizes the critical role of advanced imaging techniques and environmental control in achieving reliable and reproducible results in nanoparticle-based spectroscopic studies. Here, we verified the feasibility of this techniques to study at the single molecule levels (at nanomole concentration levels using both SERS spectroscopy and Raman imaging).

6. Summary and Conclusion

We have applied normal Raman spectroscopy, together with multivariate analysis techniques such as DAPC for differentiating Ras protein isoforms, which play significant roles in cellular signal transduction and are implicated in many cancers. A Raman library for Ras isoforms with GDP and GTP loading was built. By analyzing the amide I band, the intensity ratios of tyrosine, and Fermi doublet peaks, this study provided a detailed structural fingerprint for each isoform. We successfully predicted the secondary structure of each isoform, the hydrogen bonding condition of the phenyl side chain, and the hydrophobic condition of the tyrosine in each protein isoform. We have generated Raman barcode based on the deconvoluted Raman bands, providing the flexibility for electronic record keeping and future identification using artificial intelligence. We also highlighted the potential of Raman spectroscopy to observe the effects of specific inhibitors, like ARS-1620, on mutant KRAS G12C proteins. The ability to distinguish between similar Raman spectra of different Ras isoforms offers valuable insights into their distinct biological roles and interactions, paving the way for targeted therapeutic strategies.

The vacuum-enhanced micro-Raman spectroscopy (VERS) technique was innovatively developed to amplify weak Raman signals from low-concentration liquid samples. By utilizing vacuum evaporation to concentrate analytes within a micron-sized laser focus, VERS significantly enhances Raman signals without altering their spectra. This method is particularly advantageous for samples dissolved in solvents like DMSO, which have strong Raman backgrounds. Through systematic experiments with substances including rhodamine B, BSA, glucose, and ciprofloxacin, VERS demonstrated its efficacy and sensitivity, showing remarkable enhancement factors. This approach avoids the need for sensor materials like metal nanoparticles used in other enhancement techniques, offering a simpler, more reproducible method for analyzing various biomolecules and

chemicals. We developed this technique with potential application for Ras proteins analysis but because of limited sample availability, we could not test the Ras protein at this work. Since we tested for BSA proteins and it worked very well, we expect that this technique can be used for Ras proteins as well in future.

We also established the SERS and SERS-based Raman imaging technique in our laboratory for the study of Ras protein at nanomole concentration levels. In this work, we tested this setup with R6G molecules. We were able to obtain both SERS and SERS-based Raman image of R6G at single molecule level. The Raman imaging focused on the dynamics of single-molecule SERS, investigating how environmental conditions and laser power affect the stability and intensity of Raman signals. Using the advanced image analysis techniques like ThunderSTORM, the research identified and analyzed the behavior of hot particles under various conditions. Our findings revealed that lower laser powers maintained signal stability and reduced fluctuations, whereas higher powers decreased signal stability and increased fluctuations, possibly due to changes in the nanoparticle clusters. The study emphasized the need for precise control over experimental variables to harness the full potential of SERS, contributing to a deeper understanding of its mechanisms at the nanoscale and enhancing its application in biology and nanotechnology.

In future, we may increase the sensitivity of VERS technique by using the small diameter sample holder by overcoming the technical difficulties. Then, we can apply VERS for Ras isoforms detection. Also, we can apply the SERS and SERS-based Raman imaging technique to detect and study the Ras isoforms at single molecule level inside cell in-vivo at nanomole concentration levels.

References

[1] A. M. Rojas *et al*, "The Ras protein superfamily: evolutionary tree and role of conserved amino acids," *Journal of cell biology,* vol. 196, *(2),* pp. 189-201, 2012.

[2] D. K. Simanshu, D. V. Nissley and F. McCormick, "RAS proteins and their regulators in human disease," *Cell,* vol. 170, *(1),* pp. 17-33, 2017.

[3] S. Lu *et al*, "Drugging Ras GTPase: a comprehensive mechanistic and signaling structural view," *Chemical society reviews,* vol. 45, *(18),* pp. 4929-4952, 2016.

[4] J. Dong, J. Wen and H. Tian, "Homologs of eukaryotic Ras superfamily proteins in prokaryotes and their novel phylogenetic correlation with their eukaryotic analogs," *Gene,* vol. 396, *(1),* pp. 116-124, 2007.

[5] A. K. Kwan *et al*, "The path to the clinic: a comprehensive review on direct KRASG12C inhibitors," *Journal of experimental & clinical cancer research,* vol. 41, *(1),* pp. 27, 2022.

[6] X. Lu and K. Djabali, "Autophagic removal of farnesylated carboxy-terminal lamin peptides," *Cells,* vol. 7, *(4),* pp. 33, 2018.

[7] A. Wittinghofer and H. Waldmann, "Ras—a molecular switch involved in tumor formation," *Angewandte chemie international edition,* vol. 39, *(23),* pp. 4192-4214, 2000.

[8] I. A. Prior, F. E. Hood and J. L. Hartley, "The frequency of Ras mutations in cancer," *Cancer research,* vol. 80, *(14),* pp. 2969-2974, 2020.

[9] G. A. Hobbs, C. J. Der and K. L. Rossman, "RAS isoforms and mutations in cancer at a glance," *Journal of cell science,* vol. 129, *(7),* pp. 1287-1292, 2016.

[10] K. Chen *et al*, "Emerging strategies to target RAS signaling in human cancer therapy," *Journal of hematology & oncology,* vol. 14, *(1),* pp. 1-23, 2021.

[11] T. K. H. Lim *et al*, "KRAS G12C in advanced NSCLC: prevalence, co-mutations, and testing," *Lung cancer,* vol. 184, pp. 107293, 2023.

[12] J. M. Chalmers and P. R. Griffiths, "Handbook of vibrational spectroscopy," New York: John Wiley & Sons, 2002.

[13] R. R. Jones *et al*, "Raman techniques: fundamentals and frontiers," *Nanoscale research letters,* vol. 14, pp. 1-34, 2019.

[14] R. Pilot *et al*, "A review on surface-enhanced Raman scattering," *Biosensors,* vol. 9, *(2),* pp. 57, 2019.

[15] S. Laing *et al*, "Surface-enhanced Raman spectroscopy for in vivo biosensing," *Nature reviews chemistry,* vol. 1, *(8),* pp. 0060, 2017.

[16] B. Yeo *et al*, "Tip-enhanced Raman Spectroscopy–Its status, challenges and future directions," *Chemical physics letters,* vol. 472, *(1-3),* pp. 1-13, 2009.

[17] P. Verma, "Tip-enhanced Raman spectroscopy: technique and recent advances," *Chemical reviews,* vol. 117, *(9),* pp. 6447-6466, 2017.

[18] A. Knebl *et al*, "Fiber enhanced Raman gas spectroscopy," *TrAC Trends in analytical chemistry,* vol. 103, pp. 230-238, 2018.

[19] P. J. Eravuchira *et al*, "Hollow core photonic crystal fiber-assisted Raman spectroscopy as a tool for the detection of Alzheimer's disease biomarkers," *Journal of biomedical optics,* vol. 25, *(7),* pp. 077001, 2020.

[20] C. Domes *et al*, "Ultrasensitive detection of antiseptic antibiotics in aqueous media and human urine using deep UV resonance Raman spectroscopy," *Analytical chemistry,* vol. 89, *(18),* pp. 9997-10003, 2017.

[21] Y. Mizutani, "Time-resolved resonance Raman spectroscopy and application to studies on ultrafast protein dynamics," *Bulletin of chemical society of Japan.,* vol. 90, *(12),* pp. 1344-1371, 2017.

[22] K. Kneipp *et al*, "Single molecule detection using surface-enhanced Raman scattering (SERS)," *Physical review letters,* vol. 78, *(9),* pp. 1667, 1997.

[23] Y. Li *et al*, "Analysis of RAS protein interactions in living cells reveals a mechanism for pan-RAS depletion by membrane-targeted RAS binders," *Proceedings of the national academy of sciences,* vol. 117, *(22),* pp. 12121-12130, 2020.

[24] C. V. Raman and K. S. Krishnan, "A New Type of Secondary Radiation," *Nature,* vol. 121, *(3048),* pp. 501-502, 1928.

[25] E. Smith and G. Dent, *Modern Raman Spectroscopy: A Practical Approach,* John Wiley & Sons, England, pp. 23-70, 2004.

[26] P. Verma, "Tip-Enhanced Raman Spectroscopy: Technique and Recent Advances," *Chemical reviews,* vol. 117, *(9),* pp. 6447-6466, 2017.

[27] D. Cialla *et al*, "Surface-enhanced Raman spectroscopy (SERS): progress and trends," *Analytical and bioanalytical chemistry,* vol. 403, *(1),* pp. 27-54, 2012.

[28] B. Robert, "Resonance raman spectroscopy," *Photosynthesis research,* vol. 101, *(2),* pp. 147-155, 2009.

[29] A. M. Kelley, "Hyper-Raman scattering by molecular vibrations," *Annual review of physical chemistry,* vol. 61, pp. 41-61, 2010.

[30] Y. Cao and M. Sun, "Tip-enhanced Raman spectroscopy," *Reviews in physics,* vol. 8, pp. 100067, 2022.

[31] S. Wolf *et al,* "Highly Sensitive Detection of the Antibiotic Ciprofloxacin by Means of Fiber Enhanced Raman Spectroscopy," *Molecules,* vol. 24, *(24),* pp. 4512, 2019.

[32] M. Fleischmann, P. Hendra and A. J. McQuillan, "Raman spectra of pyridine adsorbed at a silver electrode," *Chemical physics letters, vol.* 26, (2), pp. 163-166, 1974.

[33] M. G. Albrecht and J. A. Creighton, "Anomalously intense Raman spectra of pyridine at a silver electrode," *Journal of American chemical society,* vol. 99, *(15),* pp. 5215-5217, 1977.

[34] D. L. Jeanmaire and R. P. Van Duyne, "Surface Raman spectroelectrochemistry: Part I. Heterocyclic, aromatic, and aliphatic amines adsorbed on the anodized silver electrode," *Journal of electroanalytical chemistry and interfacial electrochemistry,* vol. 84, *(1),* pp. 1-20, 1977.

[35] A. Campion and P. Kambhampati, "Surface-enhanced Raman scattering," *Chemical society reviews,* vol. 27, *(4),* pp. 241-250, 1998.

[36] G. C. Schatz, M. A. Young and R. P. Van Duyne, "Electromagnetic mechanism of SERS," Springer Berlin Heidelberg, vol. 103, pp. 19-45, 2006.

[37] K. Lin *et al,* "Size effect on SERS of gold nanorods demonstrated via single nanoparticle spectroscopy," *The journal of physical chemistry C,* vol. 120, *(37),* pp. 20806-20813, 2016.

[38] S. Nie and S. R. Emory, "Probing single molecules and single nanoparticles by surface-enhanced Raman scattering," *Science,* vol. 275, *(5303),* pp. 1102-1106, 1997.

[39] E. J. Blackie, E. C. Le Ru and P. G. Etchegoin, "Single-molecule surface-enhanced Raman spectroscopy of nonresonant molecules," *Journal of American chemical society.,* vol. 131, *(40),* pp. 14466-14472, 2009.

[40] S. Ding *et al,* "Nanostructure-based plasmon-enhanced Raman spectroscopy for surface analysis of materials," *Nature reviews materials,* vol. 1, *(6),* pp. 1-16, 2016.

[41] M. Kiguchi *et al,* "Surface enhanced Raman scattering on molecule junction," *Applied materials,* vol. 14, pp. 76-83, 2019.

[42] J. D. Spitzberg *et al,* "Plasmonic-nanopore biosensors for superior single-molecule detection," *Advanced materials,* vol. 31, *(23),* pp. 1900422, 2019.

[43] Y. Yu *et al*, "Roadmap for single-molecule surface-enhanced Raman spectroscopy," *Advanced photonics*, vol. 2, *(1)*, pp. 014002, 2020.

[44] E. C. Le Ru and P. G. Etchegoin, "Single-molecule surface-enhanced Raman spectroscopy," *Annual review of physical chemistry*, vol. 63, pp. 65-87, 2012.

[45] R. C. Maher, "SERS hot spots," *Springer materials*, Heidelberg, pp. 215-260, 2012.

[46] D. P. dos Santos *et al*, "Electrochemical control of the time-dependent intensity fluctuations in surface-enhanced Raman scattering (SERS)," *The Journal of physical chemistry C*, vol. 113, *(41)*, pp. 17737-17744, 2009.

[47] D. P. Dos Santos, M. L. Temperini and A. G. Brolo, "Intensity fluctuations in single-molecule surface-enhanced Raman scattering," *Accounts of Chemical Research*, vol. 52, *(2)*, pp. 456-464, 2019.

[48] P. A. Mosier-Boss, "Review of SERS substrates for chemical sensing", *Nanomaterials*, vol. 7, *(6)*, pp. 142, 2017.

[49] J. Turkevich, P. C. Stevenson and J. Hillier, "The size and shape factor in colloidal systems," *A general discussion of the Faraday society*, vol. 11, pp. 55, 1951.

[50] C. Jing *et al*, "Chrominance to dimension: a real-time method for measuring the size of single gold nanoparticles," *Analytical chemistry*, vol. 84, *(10)*, pp. 4284-4291, 2012.

[51] J. M. Brockman, B. P. Nelson and R. M. Corn, "Surface plasmon resonance imaging measurements of ultrathin organic films," *Annual review of physical chemistry*, vol. 51, pp. 41-63, 2000.

[52] K. L. Kelly *et al*, "The optical properties of metal nanoparticles: the influence of size, shape, and dielectric environment," *The Journal of physical chemistry B*, vol. 107, *(3)*, pp. 668-677, 2003.

[53] A. J. Haes *et al*, "Plasmonic materials for surface-enhanced sensing and spectroscopy," *MRS bulletin*, vol. 30, *(5)*, pp. 368-375, 2005.

[54] E. Le Ru and P. Etchegoin, *Principles of Surface-Enhanced Raman Spectroscopy: And Related Plasmonic Effects.*, *Elsevier*, pp. 1-663, 2009.

[55] S. Schlücker, "Surface-enhanced Raman spectroscopy: concepts and chemical applications," *Angewandte chemie international edition*, vol. 53, *(19)*, pp. 4756-4795, 2014.

[56] H. Ueba, S. Ichimura and H. Yamada, "Where are we in the study of SERS? role of chemisorption and charge transfer," *Surface science*, vol. 119, *(2-3)*, pp. 433-448, 1982.

[57] S. K. Saikin *et al,* "On the chemical bonding effects in the Raman response: Benzenethiol adsorbed on silver clusters," *Physical chemistry chemical physics,* vol. 11, *(41),* pp. 9401-9411, 2009.

[58] G. P. Smith *et al,* "Raman imaging of drug delivery systems," *Advance drug delivery reviews,* vol. 89, pp. 21-41, 2015.

[59] E. Lee, "Raman imaging: Techniques and applications," *Springer series in optical sciences,* vol. 2, pp. 85-118, 2012.

[60] Y. Shen, F. Hu and W. Min, "Raman imaging of small biomolecules," *Annual review of biophysics,* vol. 48, pp. 347-369, 2019.

[61] M. Schmick, A. Kraemer and P. I. Bastiaens, "Ras moves to stay in place," *Trends in cell biology,* vol. 25, *(4),* pp. 190-197, 2015.

[62] F. Eisenhaber, B. Persson and P. Argos, "Protein structure prediction: recognition of primary, secondary, and tertiary structural features from amino acid sequence," *Critical review in biochemistry and molecular biology,* vol. 30, *(1),* pp. 1-94, 1995.

[63] V. Lin and J. L. Koenig, "Raman studies of bovine serum albumin," *Biopolymers: Original research on biomolecules,* vol. 15, *(1),* pp. 203-218, 1976.

[64] A. Hornemann *et al,* "Intracellular SERS hybrid probes using BSA–reporter conjugates," *Analytical and bioanalytical chemistry,* vol. 405, pp. 6209-6222, 2013.

[65] D. Drescher *et al,* "SERS reveals the specific interaction of silver and gold nanoparticles with hemoglobin and red blood cell components," *Physical chemistry chemical physics,* vol. 15, *(15),* pp. 5364-5373, 2013.

[66] A. Rygula *et al,* "Raman spectroscopy of proteins: a review," *Journal of Raman spectroscopy,* vol. 44, *(8),* pp. 1061-1076, 2013.

[67] G. Zhu *et al,* "Raman spectra of amino acids and their aqueous solutions," S*pectrochimica acta part A: Molecular and biomolecular spectroscopy,* vol. 78, *(3),* pp. 1187-1195, 2011.

[68] B. Hernández *et al,* "Vibrational analysis of amino acids and short peptides in hydrated media. VIII. Amino acids with aromatic side chains: L-phenylalanine, L-tyrosine, and L-tryptophan," *The journal of physical chemistry B,* vol. 114, *(46),* pp. 15319-15330, 2010.

[69] F. Madzharova, Z. Heiner and J. Kneipp, "Surface enhanced hyper-Raman scattering of the amino acids tryptophan, histidine, phenylalanine, and tyrosine," *The journal of physical chemistry C,* vol. 121, *(2),* pp. 1235-1242, 2017.

[70] B. Fazio *et al*, "SERS detection of biomolecules at physiological pH via aggregation of gold nanorods mediated by optical forces and plasmonic heating," *Scientific reports*, vol. 6, *(1)*, pp. 26952, 2016.

[71] H. Weinmann and E. Ottow, "Recent development in novel anticancer therapies," in *Comprehensive medicinal chemistry II*, J. B. Taylor and D. J. Triggle, Eds. 2007.

[72] S. A. Forbes *et al*, "COSMIC: mining complete cancer genomes in the catalogue of somatic mutations in cancer," *Nucleic acids research*, vol. 39, pp. D945-D950, 2011.

[73] E. Castellano and E. Santos, "Functional Specificity of Ras Isoforms," *Genes & cancer*, vol. 2, *(3)*, pp. 216-231, 2011.

[74] Z. Zheng and E. C. Chang, "Chapter two - A bimolecular fluorescent complementation screen reveals complex roles of endosomes in ras-mediated signaling," in *Methods in enzymology*, P. M. Conn, Ed. 2014.

[75] J. M. Ostrem *et al*, "K-Ras (G12C) inhibitors allosterically control GTP affinity and effector interactions," *Nature*, vol. 503, *(7477)*, pp. 548-551, 2013.

[76] C. W. Johnson *et al*, "Allosteric site variants affect GTP hydrolysis on Ras," *Protein science*, vol. 32, *(10)*, pp. e4767, 2023.

[77] M. R. Janes *et al*, "Targeting KRAS mutant cancers with a covalent G12C-specific inhibitor," *Cell*, vol. 172, *(3)*, pp. 578-589. e17, 2018.

[78] A. Brockhinke *et al*, "Structural changes in the Ras protein revealed by fluorescence spectroscopy," *Physical chemistry chemical physics*, vol. 5, *(16)*, pp. 3498-3506, 2003.

[79] N. C. Maiti *et al*, "Raman spectroscopic characterization of secondary structure in natively unfolded proteins: α-synuclein," *Journal of the American chemical society*, vol. 126, *(8)*, pp. 2399-2408, 2004.

[80] J. H. Wang *et al*, "Raman difference studies of GDP and GTP binding to c-Harvey ras," *Biochemistry*, vol. 37, *(31)*, pp. 11106-11116, 1998.

[81] M. Liu *et al*, "Rapid discrimination of colon cancer cells with single base mutation in KRAS gene segment using laser tweezers Raman spectroscopy," *Journal of biophotonics*, vol. 12, *(3)*, pp. e201800332, 2019.

[82] S. Signorelli, S. Cannistraro and A. R. Bizzarri, "Raman evidence of p53-DBD disorder decrease upon interaction with the anticancer protein azurin," *International journal of molecular sciences*, vol. 20, *(12)*, pp. 3078, 2019.

[83] R. A. Copeland, S. P. Fodor and T. G. Spiro, "Surface-enhanced Raman spectra of an active flavoenzyme: glucose oxidase and riboflavin binding protein on silver particles," *Journal of American chemical society,* vol. 106, *(13),* pp. 3872-3874, 1984.

[84] M. Casella *et al*, "Raman and SERS recognition of β-carotene and haemoglobin fingerprints in human whole blood," S*pectrochimica acta part A: Molecular and biomolecular spectroscopy* vol. 79, *(5),* pp. 915-919, 2011.

[85] T. M. Cotton, S. G. Schultz and R. P. Van Duyne, "Surface-enhanced resonance Raman scattering from cytochrome c and myoglobin adsorbed on a silver electrode," *Journal of the American chemical society,* vol. 102, *(27),* pp. 7960-7962, 1980.

[86] I. R. Nabiev, V. A. Savchenko and E. S. Efremov, "Surface-enhanced Raman spectra of aromatic amino acids and proteins adsorbed by silver hydrosols," *Journal of Raman spectroscopy,* vol. 14, *(6),* pp. 375-379, 1983.

[87] X. Yang *et al*, "Highly sensitive detection of proteins and bacteria in aqueous solution using surface-enhanced Raman scattering and optical fibers," *Analytical chemistry,* vol. 83, *(15),* pp. 5888-5894, 2011.

[88] Koury, J. "Guide for Raman Shift Standards for Spectrometer Calibration in Annual Book ASTM Standards, " *ASTM: West Conshohocken*, vol. 03.06, pp. 1840, 1996.

[89] I. T. Jolliffe and J. Cadima, "Principal component analysis: a review and recent developments," *Philosophical transactions of the royal society A: Mathematical, physical and engineering sciences*, vol. 374, *(2065),* pp. 20150202, 2016.

[90] M. E. Wall, A. Rechtsteiner and L. M. Rocha, "Singular value decomposition and principal component analysis," in *A Practical Approach to Microarray Data Analysis.* Boston, MA: Springer US, pp. 91-109, 2003.

[91] T. Jombart, S. Devillard and F. Balloux, "Discriminant analysis of principal components: a new method for the analysis of genetically structured populations," *BMC genetics*, vol. 11, *(1),* pp. 94, 2010.

[92] M. Ovesný *et al*, "ThunderSTORM: a comprehensive ImageJ plug-in for PALM and STORM data analysis and super-resolution imaging," *Bioinformatics*, vol. 30, *(16),* pp. 2389-2390, 2014.

[93] S. P. Mo, J. M. Coulson, and I. A. Prior, "RAS variant signaling," *Biochemical society transactions*, vol. 46, *(5),* pp. 1325, 2018.

[94] M. Cefalì *et al*, "Research progress on KRAS mutations in colorectal cancer," *Journal of cancer metastasis and treatment,* vol. 7, pp. 26, 2021.

[95] S. P. Mulvaney and C. D. Keating, "Raman spectroscopy," *Analytical chemistry,* vol. 72, *(12),* pp. 145-158, 2000.

[96] P. Vandenabeele, *Practical Raman Spectroscopy: An Introduction.* John Wiley & Sons, 2013.

[97] R. Tuma, "Raman spectroscopy of proteins: from peptides to large assemblies," *Journal of Raman spectroscopy,* vol. 36, *(4),* pp. 307-319, 2005.

[98] Z. Wen, "Raman spectroscopy of protein pharmaceuticals," *Journal of pharmaceutical sciences,* vol. 96, *(11),* pp. 2861-2878, 2007.

[99] R. S. Senger *et al,* "Spectral characteristics of urine specimens from healthy human volunteers analyzed using Raman chemometric urinalysis (Rametrix)," *PLoS One,* vol. 14, *(9),* pp. e0222115, 2019.

[100] W. Carswell, J. L. Robertson and R. S. Senger, "Raman spectroscopic detection and quantification of macro-and microhematuria in human urine," *Applied spectroscopy,* vol. 76, *(3),* pp. 273-283, 2022.

[101] W. Liu *et al,* "Raman spectroscopy in colorectal cancer diagnostics: comparison of PCA-LDA and PLS-DA models," *Journal of spectroscopy,* vol. 20, pp. 1-6, 2016.

[102] S. Luo *et al,* "Discrimination of premalignant lesions and cancer tissues from normal gastric tissues using Raman spectroscopy," *Journal of biomedical optics,* vol. 18, *(6),* pp. 067004, 2013.

[103] I. Harada, T. Miura and H. Takeuchi, "Origin of the doublet at 1360 and 1340 cm− 1 in the Raman spectra of tryptophan and related compounds," S*pectrochimica acta part A: Molecular and biomolecular spectroscopy,* vol. 42, *(2-3),* pp. 307-312, 1986.

[104] A. Sadat and I. J. Joye, "Peak fitting applied to Fourier transform infrared and Raman spectroscopic analysis of proteins," *Applied sciences,* vol. 10, *(17),* pp. 5918, 2020.

[105] G. Pezzotti *et al,* "Raman molecular fingerprints of SARS-CoV-2 British variant and the concept of Raman barcode," *Advanced science,* vol. 9, *(3),* pp. 2103287, 2022.

[106] S. Wang *et al,* "Characterization of the dynamic germination of individual Clostridium difficile spores using Raman spectroscopy and differential interference contrast microscopy," *Journal of bacteriology,* vol. 197, *(14),* pp. 2361-2373, 2015.

[107] F. M. Lyng *et al,* "Discrimination of breast cancer from benign tumors using Raman spectroscopy," *PLoS One,* vol. 14, *(2),* pp. e0212376, 2019.

[108] M. Wu *et al,* "Stable optical trapping and sensitive characterization of nanostructures using standing-wave Raman tweezers," *Scientific reports,* vol. 7, *(1),* pp. 42930, 2017.

[109] M. Xu *et al*, "Detection of pesticide residues in food using surface-enhanced Raman spectroscopy: a review," *Journal of agricultural and food chemistry,* vol. 65, *(32),* pp. 6719-6726, 2017.

[110] B. Chazallon and C. Pirim, "Selectivity and CO_2 capture efficiency in CO_2-N_2 clathrate hydrates investigated by in-situ Raman spectroscopy," *Chemical engineering journal,* vol. 342, pp. 171-183, 2018.

[111] Y. Jiang *et al*, "Surface enhanced Raman spectroscopy (SERS): a novel reliable technique for rapid detection of common harmful chemical residues," *Trends in food science & technology,* vol. 75, pp. 10-22, 2018.

[112] B. Zachhuber *et al*, "Stand-off Raman spectroscopy: a powerful technique for qualitative and quantitative analysis of inorganic and organic compounds including explosives," *Analytical and bioanalytical chemistry,* vol. 400, *(8),* pp. 2439-2447, 2011.

[113] I. R. Lewis and H. Edwards, *Handbook of Raman Spectroscopy: From the Research Laboratory to the Process Line,* CRC press Florida, 2001.

[114] A. MYERS KELLEY, "Hyper-Raman Scattering by Molecular Vibrations," *Annual review of physical chemistry,* vol. 61, pp. 41-61, 2010.

[115] C. L. Haynes, A. D. McFarland and R. P. Van Duyne, "Surface-enhanced Raman spectroscopy," *Analytical chemistry,* vol. 77, pp. 338-346, 2005.

[116] K. G. Stamplecoskie *et al*, "Optimal size of silver nanoparticles for surface-enhanced Raman spectroscopy," *The journal of physical chemistry C,* vol. 115, *(5),* pp. 1403-1409, 2011.

[117] R. Petry, M. Schmitt and J. Popp, "Raman spectroscopy—a prospective tool in the life sciences," *Chemphyschem,* vol. 4, *(1),* pp. 14-30, 2003.

[118] S. Hanf *et al*, "Fiber-enhanced Raman multigas spectroscopy: a versatile tool for environmental gas sensing and breath analysis," *Analytical chemistry,* vol. 86, *(11),* pp. 5278-5285, 2014.

[119] P. C. Lee and D. Meisel, "Adsorption and surface-enhanced Raman of dyes on silver and gold sols," *Journal of physical chemistry,* vol. 86, *(17),* pp. 3391-3395, 1982.

[120] M. Li *et al*, "Heterostructured cube Au–Ag composites for rapid Raman detection of antibiotic ciprofloxacin," *Journal of Raman spectroscopy,* vol. 48, *(4),* pp. 525-529, 2017.

[121] I. J. Hidi *et al*, "Ciprofloxacin: pH-dependent SERS signal and its detection in spiked river water using LoC-SERS," *Analytical and bioanalytical chemistry,* vol. 408, pp. 8393-8401, 2016.

[122] A. Michałowska and A. Kudelski, "The first silver-based plasmonic nanomaterial for shell-isolated nanoparticle-enhanced Raman spectroscopy with magnetic properties," *Molecules*, vol. 27, *(10)*, pp. 3081, 2022.

Appendix A: PCA code for training and testing model

```
clear all
Train = load("TrainingData.txt");
Test = load("TestingData.txt");

%Normalization of the training data (965 X 420) the data
SquareSum=zeros(1,420);
for i=1:420
   for h=1:965
      SquareSum(1,i)=SquareSum(1,i)+Train(h,i)^2;
   end
end
SqrtSum = sqrt(SquareSum);

for i=1:420
      for h=1:965
       Training(h,i)=Train(h,i)/SqrtSum(1,i);
   end
end
%Normalization of the training data (965 X 30) the data
SquareSum=zeros(1,30);
for i=1:30
   for h=1:965
      SquareSum(1,i)=SquareSum(1,i)+Test(h,i)^2;
   end
end
SqrtSum = sqrt(SquareSum);

for i=1:30
      for h=1:965
       Testing(h,i)=Test(h,i)/SqrtSum(1,i);
   end
end

 %PCA of the training data
[coeff,score,latent] = pca(Training);
%Normalization of score
SquareSum=zeros(1,420);
for i=1:420
   for h=1:965
      SquareSum(1,i)=SquareSum(1,i)+score(h,i)^2;
   end
end
SqrtSum = sqrt(SquareSum);
for i=1:420
      for h=1:965
       score1(h,i)=score(h,i)/SqrtSum(1,i);
   end
end
    PC = (Training - mean(Training))'*score1;    % coeff of training data
    PC1 = (Testing - mean(Testing))'*score1;   % coeff of test data
  hold on;
```

```
  %plotting testing data
scatter3(PC(1:28,1),PC(1:28,2),PC(1:28,3),80,'o','r');
scatter3(PC(29:56,1),PC(29:56,2),PC(29:56,3),80,'v','r');
scatter3(PC(57:84,1),PC(57:84,2),PC(57:84,3),80,'h','r');

scatter3(PC(85:112,1),PC(85:112,2),PC(85:112,3),80,'o','g');
scatter3(PC(113:140,1),PC(113:140,2),PC(113:140,3),80,'v','g');
scatter3(PC(141:168,1),PC(141:168,2),PC(141:168,3),80,'h','g');

scatter3(PC(169:196,1),PC(169:196,2),PC(169:196,3),80,'o','b');
scatter3(PC(197:224,1),PC(197:224,2),PC(197:224,3),80,'v','b');
scatter3(PC(225:252,1),PC(225:252,2),PC(225:252,3),80,'h','b');

scatter3(PC(253:280,1),PC(253:280,2),PC(253:280,3),80,'o','c');
scatter3(PC(281:308,1),PC(281:308,2),PC(281:308,3),80,'v','c');
scatter3(PC(309:336,1),PC(309:336,2),PC(309:336,3),80,'h','c');

scatter3(PC(337:364,1),PC(337:364,2),PC(337:364,3),80,'o','m');
scatter3(PC(365:392,1),PC(365:392,2),PC(365:392,3),80,'v','m');
scatter3(PC(393:420,1),PC(393:420,2),PC(393:420,3),80,'h','m');

%projecting testing data to the training data plot

scatter3(PC1(1:2,1),PC1(1:2,2),PC1(1:2,3),80,'v','filled','r');
scatter3(PC1(3:4,1),PC1(3:4,2),PC1(3:4,3),80,'o','filled','r');
scatter3(PC1(5:6,1),PC1(5:6,2),PC1(5:6,3),80,'h','filled','r');

scatter3(PC1(7:8,1),PC1(7:8,2),PC1(7:8,3),80,'v','filled','g');
scatter3(PC1(9:10,1),PC1(9:10,2),PC1(9:10,3),80,'o','filled','g');
scatter3(PC1(11:12,1),PC1(11:12,2),PC1(11:12,3),80,'h','filled','g');

scatter3(PC1(13:14,1),PC1(13:14,2),PC1(12:13,3),80,'v','filled','b');
scatter3(PC1(15:16,1),PC1(15:16,2),PC1(15:16,3),80,'o','filled','b');
scatter3(PC1(17:18,1),PC1(17:18,2),PC1(17:18,3),80,'h','filled','b');

scatter3(PC1(19:20,1),PC1(19:20,2),PC1(19:20,3),80,'v','filled','c');
scatter3(PC1(21:22,1),PC1(21:22,2),PC1(21:22,3),80,'o','filled','c');
scatter3(PC1(23:24,1),PC1(23:24,2),PC1(23:24,3),80,'h','filled','c');

scatter3(PC1(25:26,1),PC1(25:26,2),PC1(25:26,3),80,'v','filled','m');
scatter3(PC1(27:28,1),PC1(27:28,2),PC1(27:28,3),80,'o','filled','m');
scatter3(PC1(29:30,1),PC1(29:30,2),PC1(29:30,3),80,'h','filled','m');

xlabel('PC 1');
ylabel('PC 2');
zlabel('PC 3');
title('PCA Results');
rotate3d on;
hold off;
```

Appendix B: DAPC code for training and testing model

```
data = load("TrainingPC.txt"); % PC (coefficient of training data)
Traindata = load("TestingPC.txt"); % PC1 (coefficient of testing data)
group =
['A';'A';'A';'A';'A';'A';'A';'A';'A';'A';'A';'A';'A';'A';'A';'A';'A';'A';'A';'A';'A';'A';'A';'A';'A';'A';'A';'B';'B';'B';'B';'B';'B';'
B';'B';'B';'B';'B';'B';'B';'B';'B';'B';'B';'B';'B';'B';'B';'B';'B';'B';'B';'B';'B';'B';'C';'C';'C';'C';'C';'C';'C';'C';'C';'C';'C';'C';'C';'C';'C';'C'
;'C';'C';'C';'C';'C';'C';'C';'C';'C';'C';'C';'C';'C';'C';'C';'D';'D';'D';'D';'D';'D';'D';'D';'D';'D';'D';'D';'D';'D';'D';'D';'D';'D';'D';'D';'D';'D
';'D';'D';'D';'D';'D';'D';'E';'E';'E';'E';'E';'E';'E';'E';'E';'E';'E';'E';'E';'E';'E';'E';'E';'E';'E';'E';'E';'E';'E';'E';'E';'E';'E';'E';'E';'E'];

[~,~,statsTrain] = manova1(data(:, 1:11),group);
 data1 = data(:,1:11);
 data2 = Traindata(:, 1:11);

C = (data1 - mean(data1))*statsTrain.eigenvec;
testC = (data2- mean(data2))*statsTrain.eigenvec;
 hold on;

% Plot 3D fot training
scatter3(C(1:28,1),C(1:28,2),C(1:28,3),30,'o','r');
scatter3(C(29:56,1),C(29:56,2),C(29:56,3),30,'v','g');
scatter3(C(57:84,1),C(57:84,2),C(57:84,3),30,'>','b');
scatter3(C(85:112,1),C(85:112,2),C(85:112,3),30,'o','c');
scatter3(C(113:140,1),C(113:140,2),C(113:140,3),30,'v','m');

   % Plot 3D fot testing
scatter3(testC(1:2,1),testC(1:2,2),testC(1:2,3),80,'o','filled','r');
scatter3(testC(3:4,1),testC(3:4,2),testC(3:4,3),80,'v','filled','g');
scatter3(testC(5:6,1),testC(5:6,2),testC(5:6,3),80,'>','filled','b');
scatter3(testC(7:8,1),testC(7:8,2),testC(7:8,3),80,'o','filled','c');
scatter3(testC(9:10,1),testC(9:10,2),testC(9:10,3),80,'v','filled','m');

xlabel('Canonical 1');
ylabel('Canonical 2');
zlabel('Canonical 3');
title('DAPC (GDP)');
rotate3d on;
hold off;
```