

**Raman Spectroscopy for Biomedical Applications:
Studies of Lysozyme Aggregation, Detection of
Extracellular Vesicles under Autophagic
Conditions and Indels in Plants**

A Thesis
Submitted for the Degree of

Doctor of Philosophy

by
DIVYA CHALAPATHI



Chemistry and Physics of Materials Unit

Jawaharlal Nehru Centre for Advanced Scientific Research
(A Deemed University)
Bangalore – 560064

March 2022

To,

My Parents and Brother

DECLARATION

I hereby declare that the matter embodied in this thesis entitled “**Raman Spectroscopy for Biomedical Applications: Studies of Lysozyme Aggregation, Detection of Extracellular Vesicles under Autophagic Conditions and Indels in Plants**” is the result of the research carried out by me at the Chemistry and Physics of Materials Unit, Jawaharlal Nehru Centre for Advanced Scientific Research, Bangalore, India, under the supervision of **Prof. Chandrabhas Narayana** and it has not been submitted elsewhere for the award of any degree or diploma.

In keeping with the general practices of reporting scientific observation, due acknowledgements have been made whenever the work described is based on the findings of other investigators. Any omission which might have occurred by oversight or error in judgement is regretted.

Bengaluru

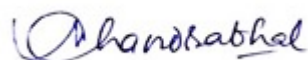


.....
Divya Chalapathi

CERTIFICATE

I hereby certify that the matter embodied in this thesis entitled “**Raman Spectroscopy for Biomedical Applications: Studies of Lysozyme Aggregation, Detection of Extracellular Vesicles under Autophagic Conditions and Indels in Plants**” has been carried out by **Ms. Divya Chalapathi** at the Chemistry and Physics of Materials Unit, Jawaharlal Nehru Centre for Advanced Scientific Research, Bangalore, India under my supervision and that has not been submitted elsewhere for the award of any degree or diploma.

Bengaluru



Professor Chandrabhas Narayana

(Research Supervisor)

ACKNOWLEDGEMENTS

I take this opportunity to acknowledge and express my gratitude to all the people who have inspired, motivated, and supported me throughout my Ph.D. journey.

I shall begin by thanking my research supervisor, Prof. Chandrabhas Narayana, for giving me an opportunity to work with his group and for guiding me through the course of my Ph.D. Being an expert in Raman spectroscopy, he has not shied away from researching broad fields of science. This enabled him to build an interdisciplinary group that works in various fields such as material science, nanotechnology, chemistry, and biology. As a result, it gave me the flexibility to pursue my interests with the additional perks of gaining insights from group members who work on a variety of topics. His connections in the scientific world aided me in landing some really fascinating collaborative projects, allowing me to investigate the strength of Raman spectroscopy from a multidimensional perspective. He encouraged us to pursue all of our ideas with complete scientific and financial backing. I sincerely thank him for his support, extensive freedom, fruitful discussions, feedback, and belief in me as a research scholar.

I want to express my heartfelt gratitude to Prof. Rajaram Swaminathan and Dr. Amrendra Kumar from the Department of bioscience and bioengineering, IIT–Guwahati; Dr. Shijulal Nelson Sathi and Mr. Pratik Behera from Rajiv Gandhi Centre for Biotechnology; Prof. Ravi Manijthaya and Dr. Sreedevi from Autophagy lab, MBGU, JNCASR; Dr. Jayaprakash from Biodecipher, Bengaluru; and Dr. Neha Vyas from St.John’s Hospital, Bengaluru, for fruitful collaborations. Through these diverse projects, I have been able to investigate and enhance my knowledge and further learn about their perspectives and approaches of the same.

I would like to thank all my course instructors, Prof. A. Sundaresan, Prof. U. V. Waghmare, Prof. S. M. Shivaprasad and Prof. Shobhana Narasimhan from

JNCASR, Bengaluru; and Prof. T. N. Guru Row from IISc, Bengaluru, for their informative courses and fruitful discussions.

I thank Prof. C.N.R. Rao and all the past and present chairpersons of CPMU for providing us with extensive research facilities and scientific temper through various talks and the Unit days. I would like to highlight that JNCASR has some of the top experimental facilities in the nation, which has greatly aided me in carrying out all of my experiments rapidly and effectively. I want to thank Prof. A. Sundaresan, in particular, for all his assistance with everyday and official lab activities when Prof. Chandrabhas was on lien.

While knowledge and hard work are integral, financial assistance is also crucial. Hence, I thank the DST-INSPIRE and Sheikh Saqr fellowship for financial aid.

I would like to thank all CPMU lab technicians, especially Vasu sir, Usha madam, Mr. Rahul and Srinath sir. I would like to recognise and thank the CPMU office staff, Pavana, Radha, Usha and Balraj, for their efforts to date. Furthermore, I would like to thank all of the many sections of JNCASR, including academic, administrative, complab, library, stores and purchase, DG room, audio-visual team, security, dhanvantri, hostel, mess, dining hall and utility, which are all necessary for the efficient completion of one's Ph.D.

I would like to thank all the former members of the Light Scattering Lab, Dr. Gayatri, Dr. Partha, Dr. Dhanya, Dr. Sayan, Dr. Rajaji, Dr. Kamali, Priyank, Dr. Shantanu, Sunami, Samyabrata Sen and Suhas, for their informative discussions and a pleasant atmosphere in the lab, during my Ph.D. I would want to highlight that working closely with Samyabrata and Sunami on different projects was a fantastic scientific experience.

No words are enough to describe my present labmates, Priyanka, Janaky and Anjana. They have become very dear to me and have been really kind, joyous, and supportive. They are everything and more than I could ask for at every stage of my

Ph.D. in this lab. Each of their own bright individualities has shaped me to be the person I am today, and I shall be immensely grateful for that.

While I was mostly preoccupied with my Ph.D. research, time spent at ETU with the outreach activities was always a refreshing break. It was a nice experience working with Vinayak, Praveen and all the SMP volunteers. I would like to express my gratitude to Dr. Indumati Rao for providing me with freedom and this fantastic opportunity to explore scientific teaching to high school kids. I always looked forward to all the special moments and the delicious snacks and goodies provided by her.

I would like to thank Shobha Madam and Sushmita for their affection, hospitality and the many fun times we had together.

I express my thanks to Prof. N.S. Vidhyadhiraja and Jayachandra sir for their mentoring on several occasions outside of lab activities.

I would like to thank my JNCASR buddies, Pradeep, Vybhav, Priyanka, Premakumar, Shivaram, Srinath, Manaswee, Ravi, Deepak, Bharath and Rajendra who have become my family away from home. The countless memories that I have built with them, from everyday activities to sports, movies, dinners, and trips will always have a special place in my heart. I would also like to thank Anusha, Malay, Usha, Swati, Vandhana, Ganesh, Asutosh, Shreya, Nimish, Shashi, Priya, Pavitra, Rafi, Vijay, Sourav, Shrilaxmi and Suchi, for making the journey more memorable.

I would like to thank my friends from outside of JNCASR, especially Prasiddha, Kamala and Teerna, for their motivation and unconditional support throughout.

I also thank my mama's and aunt's families for the great time I had over many weekends with them.

Finally, I would like to express my sincere gratitude to my parents, G.S.V Chalapathi and Suchetha Chalapathi, and my brother, Darshan, for always believing in me, and giving me their unconditional love, strength and support.

PREFACE

This thesis aims at exploiting Raman Spectroscopy for various biomedical applications. It highlights the tool's versatility ranging from understanding microenvironmental effects on proteins during their aggregation, using Surface Enhanced Raman Spectroscopy (SERS) to distinguish Extracellular Vesicles based on their origin, all the way to designing a strategy to detect Indels in tomato plants. Along with these, a low-cost tabletop Raman spectrometer with automated XYZ mapping was constructed that can find applications in various settings.

Chapter 1 introduces the basic concepts of Raman Spectroscopy, SERS, Raman spectroscopy of biomolecules and their applications for biomolecular detection.

Chapter 2 describes the various experimental and characterisation techniques that were used throughout the thesis.

Chapter 3 demonstrates the application of Raman Spectroscopy in understanding the Lysozyme protein aggregation in alkaline pH. The Raman spectra gave multiple insights into the effects of aggregation compared to many other techniques that provide one or few properties per study. The shift in the Amide I band from 1657 cm^{-1} to 1672 cm^{-1} reveals that upon incubation in a pH 12.2 buffer at $25\text{ }^{\circ}\text{C}$, hen egg-white lysozyme (HEWL) goes from an alpha-helical dominated system to a beta-sheet dominated structure, in ten days. The inhibitor molecule studies show that disulphide links are essential for polymerization, resulting in beta-sheet structures.

Chapter 4 illustrates the potential of combining Principal Component Analysis with SERS to detect Extracellular vesicles (EVs) laden with lipid, protein, DNA, micro-RNAs that play essential roles in intercellular communication and waste disposal. Simple citrate reduced silver nanoparticles assisted SERS was used to distinguish EVs extracted from several cell lines isolated under normal and autophagic conditions (nitrogen starvation). This study is the first report of its kind in

characterising EVs from cells under autophagic conditions using SERS. We used two cancerous cell lines, HeLa, its corresponding autophagy-deficient cell line (Atg5^{-/-}), and a non-cancerous cell line, HEK293, to isolate EVs. Along with the differentiation of EVs, few biochemical insights were obtained through these studies.

Chapter 5 involves the development of a PCR-based SERS strategy to detect indels in tomato plants. The Tomato Yellow Leaf Curl (TYLC) virus transmitted by the Silverleaf Whitefly causes its leaves to turn yellow, stunt its growth, reduce fruit production and eventually cause the death of the plant. A presence of an Indel in the Ty3 gene makes the plant resistant to the attack of this virus. With this as inspiration, we developed a strategy that involves the design of an oligomeric probe with a Raman active dye molecule that can help predict the presence or absence of the indel in the extracted DNA sample.

Chapter 6 consists of constructing a miniature tabletop Raman Spectrometer that can find applications ranging from biomedical detection, material characterisation, pharmaceuticals to educational purposes. A low-cost benchtop Raman spectrometer was assembled, calibrated and demonstrated. The XYZ mapping stage was automated to accommodate a 96 well plate. It was compared with a commercially available research-grade spectrometer and was seen to have good sensitivity, functionality and sturdiness.

Chapter 7 summarizes the thesis and provides an outlook to the various chapters discussed above.

PUBLICATIONS

Publications in this thesis:

- 1) *Surface-Enhanced Raman Spectroscopy as a Tool for Distinguishing Extracellular Vesicles under Autophagic Conditions: A Marker for Disease Diagnostics*

Divya Chalapathi, Sreedevi Padmanabhan, Ravi Manjithaya and Chandrabhas Narayana

J. Phys. Chem. B, 124, 48, 10952 (2020)

- 2) *PCR-based strategy for detecting mutations in tomato plants using Surface Enhanced Raman Spectroscopy*

Divya Chalapathi, Samyabrata Sen, Jayaprakash Targolli and Chandrabhas Narayana

(Manuscript under preparation)

- 3) *Understanding aggregation of Hen Egg White Lysozyme in alkaline pH using Raman Spectroscopy and MD Simulations*

Divya Chalapathi, Amrendra Kumar, Pratik Behera, Shijulal Nelson Sathi, Rajaram Swaminathan and Chandrabhas Narayana

(Manuscript under preparation)

Miscellaneous publications:

- 1) *Phase Transitions in Metals (Book Chapter)*

Divya Chalapathi, Priyanka Jain and Chandrabhas Narayana

Advances in the chemistry and Physics of Materials, World Scientific, 249-274 (2019)

- 2) *Single step synthesis of reduced graphene oxide/SnO₂ nanocomposites for potential optical and semiconductor applications*

A. Malik, I.Roy, **D.Chalapathi**, C.Narayana, T.D.Das, A.Bhattacharya, S.Bera, S.Bhattacharya, Sriparna De, B. Das, D. Chattopadhyay

Materials Science and Engineering B, 264, 114938 (2021)

Table of Contents

Chapter 1.	Introduction	1
1.1.	Raman Scattering.....	2
1.1.1.	Classical Picture of Raman Scattering	4
1.1.2.	Quantum Picture of Raman Scattering.....	6
1.2.	Information from a Raman spectrum	8
1.3.	Surface-enhanced Raman spectroscopy.....	10
1.3.1.	Factors affecting SERS enhancement.....	11
1.3.1.1.	Electromagnetic enhancement	11
1.3.1.2.	Chemical Enhancement	15
1.3.2.	Nanostructures for SERS	16
1.4.	Raman spectroscopy for biosystem analysis and detection	18
1.5.	Scope of this thesis	26
1.6.	References	27
Chapter 2.	Experimental Section	32
2.1.	UV-Visible Absorption	32
2.2.	Dynamic Light Scattering and Zeta Potential.....	33
2.3.	Transmission Electron Microscopy (TEM).....	33
2.4.	Raman Spectroscopy and SERS studies	34
2.5.	Polymerase Chain Reaction.....	36
2.6.	Synthesis of Silver Nanoparticles.....	37
2.7.	Synthesis of Gold Nanoparticles	38

Chapter 3.	Understanding aggregation of Hen Egg White Lysozyme and Human Lysozyme in alkaline pH using Raman Spectroscopy and MD simulations...	40
3.1.	Introduction	41
3.1.1.	Aggregation in HEWL	42
3.1.2.	Raman spectroscopy of proteins.....	43
3.1.3.	MD simulations for protein aggregation studies.....	47
3.2.	Materials and Methods.....	48
3.2.1.	Materials	48
3.2.2.	Experimental Methods	48
3.3.	Results and Discussion	51
3.3.1.	Unfolding of HEWL protein.....	55
3.3.2.	Protein aggregation and interactions	58
3.3.2.1.	Amide band analysis.....	59
3.3.2.2.	Effect of iodoacetamide on the aggregation	63
3.3.2.3.	Other Raman modes analysis	64
3.3.3.	Protein aggregation and interactions – MD simulations	68
3.4.	Aggregation of Human Lysozyme in alkaline pH.....	71
3.4.1.	Comparing the Raman spectra of HEWL and HL	72
3.4.2.	Amide I band analysis of HL aggregation	74
3.5.	Conclusion.....	75
3.6.	References	77
Chapter 4.	Using Surface-Enhanced Raman Spectroscopy to Distinguish Extracellular Vesicles under Autophagic Conditions	81
4.1.	Introduction	82

4.2.	Materials and Methods.....	86
4.3.	Results and Discussion	88
4.3.1.	Western Blot and TEM characterisation	88
4.3.2.	Surface-Enhanced Raman Spectroscopy	89
4.3.2.1.	Biochemical analysis of the SERS spectra	93
4.3.3.	Principal Component Analysis	96
4.3.4.	SERS for EVs extracted by ultracentrifugation method	99
4.4.	Conclusions	102
4.5.	References	104
Chapter 5.	PCR-based strategy for detecting mutations in tomato plants using Surface Enhanced Raman Spectroscopy	110
5.1.	Introduction	111
5.1.1.	Mutations – Indels and SNPs.....	112
5.1.2.	PCR based strategies to detect mutations	113
5.1.3.	PCR-based Raman spectroscopy to detect mutations.....	113
5.1.4.	Designing oligonucleotide sequences for PCR-based SERS	114
5.1.4.1.	Primer design parameters	114
5.1.4.2.	Design parameters for DNA probes	115
5.1.4.3.	Raman active dye molecules for SERS	117
5.2.	Materials and Methods.....	118
5.3.	Results and Discussions.....	122
5.3.1.	PCR-based SERS strategy	122
5.3.2.	Detection of Ty-3 indel by PCR-based SERS	125
5.3.3.	Multiplexing studies using PCR-based SERS.....	128

5.3.4.	Detecting Mi-1 indel using PCR-based SERS	130
5.3.5.	Using higher concentration ST in PCR-based SERS.....	131
5.3.6.	Detecting SNP in Ty-3 gene using PCR-based SERS	134
5.4.	Conclusions	136
5.5.	References	137
Chapter 6.	Construction of a miniaturised Raman spectrometer for applications in bio-detection.....	140
6.1.	Introduction	141
6.2.	Essential components of a Raman spectrometer.....	142
6.3.	Construction of the miniature Raman spectrometer.....	151
6.4.	Conclusion.....	163
6.5.	References	164
Chapter 7.	Summary and Outlook	165

Chapter 1. Introduction

Transmission, reflection, absorption, elastic and inelastic scattering constitute the light-matter interactions. The behaviour of these various phenomena can be exploited to get a better understanding of the material properties. For example, the UV-visible absorption spectroscopy measures the attenuation of the light in the wavelengths ranging from ultraviolet (UV), visible to near-infrared (IR) after it passes through the sample.¹ This technique can help obtain quantitative measurements of material components and also characterise optical and electronic properties of a material using this spectroscopy. Similarly, the scattering of light from different materials can be studied to obtain their several physical and chemical properties.

In the 1860s, John Tyndall made a fascinating observation when he shone light on a tube that was being gradually filled with smoke. The beam appeared blue from the sides, but it seemed to be red from the far end. The scattering of light by particles in a colloid or a fine suspension was then called the Tyndall effect. This was the first step towards explaining the blue colour of the sky. In the 1870s, James Maxwell postulated that when light scatters off a particle, the scattered radiation carries information about its molecular properties. A while after this, Lord Rayleigh theoretically analysed and established the scattering of light from suspended particles.² The Rayleigh scattering law expresses that the amount of scattering of light is conversely related to the fourth power of the incident wavelength. The origin of the colour of the sky was then explained to be the scattering of light from macroparticles suspended like dust, water droplets and ice particles.

In the 1920s, Adolf Smekal established the existence of inelastic scattering of light from particles through his theoretical studies.³ Sir C.V. Raman was able to attest that the blue colour of the sea is a result of a mix of the absorption and the scattering of sunlight by the water molecules and not just the reflection of the sky in the sea. He, along with K.S. Krishnan, was able to show that when a sample was irradiated with the violet light of the spectrum, along with violet, it also scattered light of a different

wavelength that was isolated using a green filter.^{4, 5} He was awarded the Nobel Prize in 1930 for his discovery which was later known as the “Raman Effect”. He then methodically developed a spectroscope that contained a piece of quartz that could photograph the spectrum of the scattered light and measure its wavelength.

Around this time in history, many spectroscopic tools were being developed to exploit the various aspects of molecular transitions. Depending on the energy scales probed, molecular spectroscopy can be electronic, vibrational, or rotational. Electronic spectroscopy can be studied using UV or visible radiation, vibrational spectroscopy using IR radiation or inelastic Raman scattering and rotational spectroscopy using microwave radiation. As every molecule has a characteristic vibrational spectrum, which can be used as their fingerprints, vibrational spectroscopy has become the most extensively used tool. Over years of study, it has been proved effective to study and understand the structure and dynamics of molecules and thus advantageous in molecular detection, sensing, and characterisation. Vibrational spectroscopic techniques find applications in various fields, including physical sciences, chemistry, pharmacy, cosmetics, food, medicine, agriculture, and astronomy. Due to its wide outreach, extensive research has been carried out over the past century to make these techniques stronger, more sensitive and provide more detailed molecular information. This has been achievable by advances in various fields like optical instrumentation, nanoscience, and biotechnology. The invention of lasers that can produce monochromatic light with high spatial and temporal coherence has led to the increased sensitivity and higher resolution of these techniques, hence providing a better microscopic picture. This thesis focuses on the use of a molecular spectroscopy technique, the Raman spectroscopy and its variants, to study the chemical properties of various biomaterials and also to detect them.

1.1. Raman Scattering

When light is incident on the matter, the molecules can get excited from a stable ground state to a higher energy state through several different pathways. After

excitation, the molecule transitions back to the ground state by emitting a photon. Most of the emitted photons have the same energy, while a few may lose or gain energy after interaction with the molecules. The elastic scattering where the energy of the photons doesn't change is called the Rayleigh Scattering. The most common inelastic scattering where the energy of the photons is different from the incident photon is called the Raman scattering. The Raman scattering is either anti-Stokes or Stokes scattering, depending on whether the photon has gained or lost energy (Fig. 1.1).

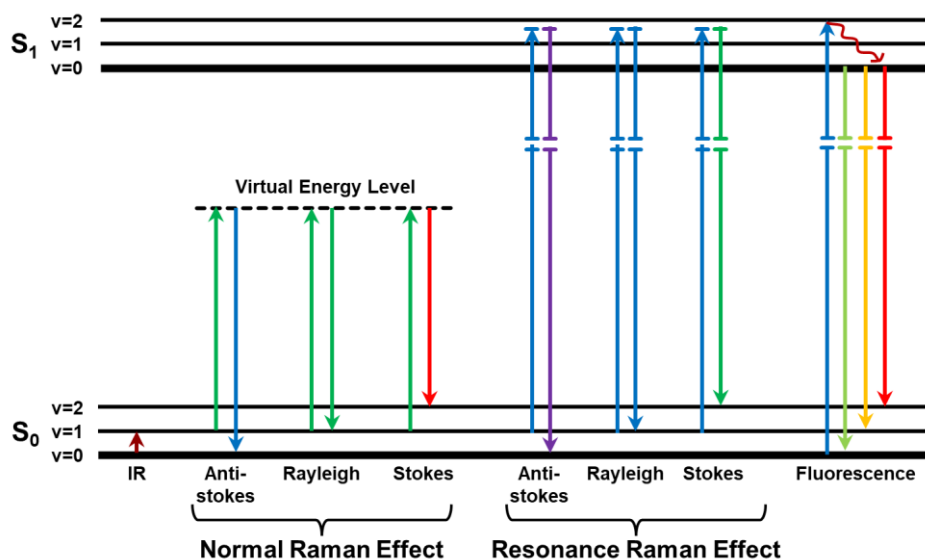


Fig. 1.1. Energy level diagram comparing IR, Rayleigh scattering, Raman scattering, Resonance Raman scattering and fluorescence.

Raman and IR absorption spectroscopy are two vibrational spectroscopy tools. In IR absorption spectroscopy, when an IR photon is incident, the molecule transitions to a higher vibrational energy level within the same electronic state, thus absorbing the photon. In comparison, Raman scattering is a two-photon process (Fig. 1.2).

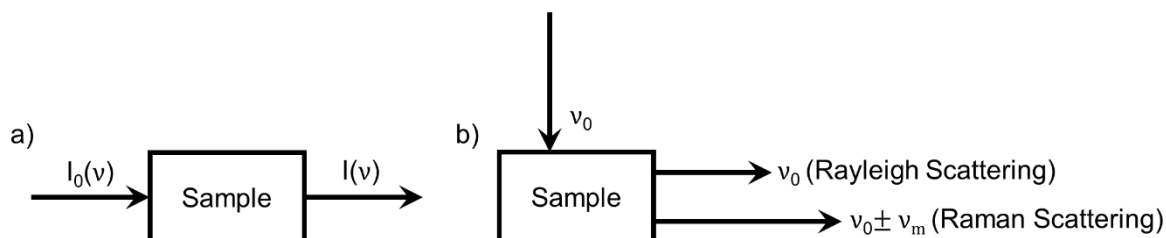


Fig. 1.2. Differences in the mechanism of a) IR absorption vs b) Raman scattering.

The incoming radiation induces a dipole in the molecule depending on the polarizability. The induced dipole can oscillate with the same frequency as the exciting

radiation or with an altered frequency due to the molecule's vibrations. Thus the scattered light has both Rayleigh and Raman scattered photons. The polarizability tensor has two cartesian components, one for the incident photon and the other for the scattered photon. A quantum mechanical process causes a single coherent event to connects the two photons. Rayleigh and Raman scattering are thus two-photon processes involving the absorption of one followed by the emission of a different photon. Raman scattering can consist of vibrational and rotational transitions but never be purely rotational. The following part of this chapter discusses the classical and quantum pictures of Raman scattering.

1.1.1. Classical Picture of Raman Scattering

The classical approach to Raman scattering is based on the first-order of induced electrical dipoles. In a static electric field, \vec{E} the positively charged nucleus and the negatively charged electron cloud of a molecule get spatially distorted. This separation of charges creates an induced dipole moment, \vec{p} and the molecule is polarized. The strength of the induced polarization is directly proportional to the applied external electric field. The polarization is given by the following equation,

$$\vec{P} = \hat{\alpha} \cdot \vec{E} \quad (1.1)$$

where $\hat{\alpha}$ is the polarizability of the molecule represented by a 3×3 polarizability tensor, consisting of 9 components in the cartesian coordinate system. The tensor components depend on the positions of the nuclei of the atoms with respect to the electron cloud. They thus are sensitive to the changes in the bonding hence the molecular coordinates. The polarization equation can be represented as,

$$\begin{bmatrix} P_x \\ P_y \\ P_z \end{bmatrix} = \begin{bmatrix} \alpha_{xx} & \alpha_{xy} & \alpha_{xz} \\ \alpha_{yx} & \alpha_{yy} & \alpha_{yz} \\ \alpha_{zx} & \alpha_{zy} & \alpha_{zz} \end{bmatrix} \begin{bmatrix} E_x \\ E_y \\ E_z \end{bmatrix} \quad (1.2)$$

When a molecule is exposed to a time-varying electric field of frequency ν ,

$$E = E_0 \cos 2\pi\nu t \quad (1.3)$$

then the magnitude of the induced dipole moment varies as,

$$P = \hat{\alpha} \cdot \vec{E} = \alpha E_0 \cos 2\pi\nu t \quad (1.4)$$

This oscillatory dipole produces Rayleigh scattering, the elastic scattering component. Molecules have internal degrees of freedom like vibration or rotation that periodically modulate the polarizability of the molecule. Hence, the oscillating dipole is a superposition of the varying electric field and the polarizability of the molecule. A molecule consisting of N atoms has $(3N-6)$ vibrational modes if it is non-linear and $(3N-5)$ modes if it is linear. The polarizability of a molecule is a function of its independent normal vibrational modes. If the molecule vibrates with a normal mode of frequency ν_m , the nuclear displacement q is given by,

$$q = q_0 \cos 2\pi\nu_m t \quad (1.5)$$

where q_0 is the amplitude of the vibration. Thus, for a small amplitude of vibration, the polarizability can be expressed as,

$$\alpha = \alpha_0 + \left(\frac{\partial \alpha}{\partial q} \right)_{q=0} q + \dots \quad (1.6)$$

where α_0 is the equilibrium polarizability and $\left(\frac{\partial \alpha}{\partial q} \right)_{q=0}$ is the rate of change of polarizability. Then the magnitude of the induced dipole moment is given by,

$$P = \alpha E_0 \cos 2\pi\nu t + \frac{1}{2} \left(\frac{\partial\alpha}{\partial q} \right)_{q=0} q_0 E_0 [\cos 2\pi(\nu + \nu_m) - \cos 2\pi(\nu - \nu_m)] \quad (1.7)$$

From this equation, we can see that the oscillating dipole has three frequency terms, ν , $\nu + \nu_m$ and $\nu - \nu_m$. The component ν corresponds to the elastic Rayleigh scattering, where the frequency of the scattered light is the same as the incident light. The terms, $\nu + \nu_m$ and $\nu - \nu_m$ correspond to the inelastically scattered components, i.e., the Stokes and anti-Stokes scattered radiation, respectively. If the vibrational mode doesn't alter the polarizability of the molecule, then $\left(\frac{\partial\alpha}{\partial q} \right)_{q=0} = 0$, and the frequency of the induced dipole moment is not altered. Thus, for a vibrational mode to be active, the molecular polarizability needs to change, and it should be a monotonic function of the molecular displacement.

1.1.2. Quantum Picture of Raman Scattering

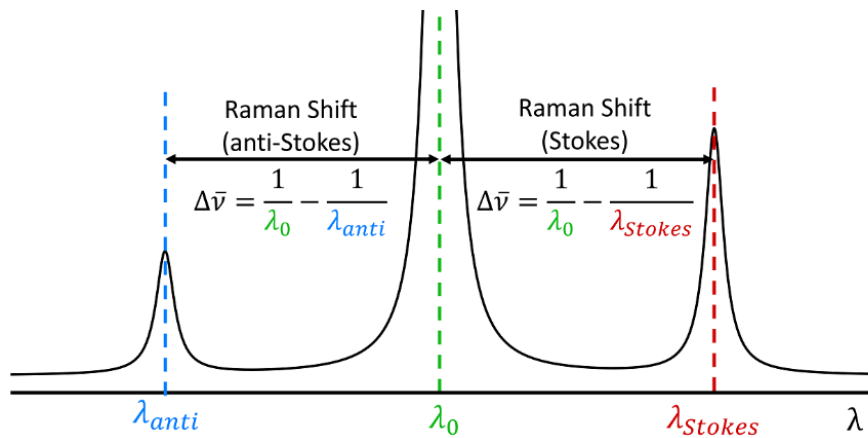


Fig. 1.3. Schematic representation of the intensities of the Stokes and anti-Stokes Raman peaks.

The origin of the Stokes and anti-Stokes lines can be well understood from the classical theory of Raman scattering. According to this classical picture, both Stokes and anti-Stokes scattering lines have the same intensity. But it is always observed that Stokes scattering that involves loss of energy of the photon gives a more intense signal than anti-Stokes scattering and the intensity ratio of anti-Stokes to Stokes keeps decreasing with increasing wavenumber (Fig. 1.3). Hence, the quantum mechanical picture is needed to get a better understanding of this difference in intensities. This

quantum picture of Raman scattering was derived by Placzek⁶ using the dispersion theory given by Kramers and Heisenberg.^{7, 8} It is a two-photon process based on second-order perturbation theory. In this picture, the molecule is treated quantum mechanically, while the radiation is treated classically.

According to the fundamentals of quantum mechanics, the energy corresponding to the electronic, vibrational, and rotational degrees of freedom of a molecule, can take discrete values. This discrete set consists of quantized energy levels corresponding to the possible stationary states of the molecule. These states are characterised by a specific set of quantum numbers describing the level of excitation of each quantum degree of freedom and by a corresponding wave function. A molecule can transition from one vibrational quantum state of energy E_1 (S_0) to another state having energy E_2 (S_1) through a virtual state having energy $E\nu$ (Fig. 1.1). The virtual state is an imaginary state that doesn't have a corresponding eigenstate of the molecule. The incident electromagnetic radiation perturbs the molecule. If the transition's dipole moment is non-zero, a direct transition between two energy levels is accompanied by emission or absorption. The amplitude for a transition from an initial state i to a final state f induced by radiation of wavenumber $\tilde{\nu}_0$ is,

$$[\mu]_{fi} = \langle \psi_f | \alpha | \psi_i \rangle \cdot E \quad (1.8)$$

where ψ_f and ψ_i are the wavefunctions of the final and initial states. The transition polarizability tensor determines the intensity of Raman scattered radiation. By ignoring the electrical and mechanical anharmonicity only for the non-zero elements, the vibrational quantum number changes by unity, whereas all other quantum numbers remain unchanged. The typical matrix element of the transition polarizability associated with Stokes Raman scattering at Raman shift is,

$$[\alpha_{xy}] = b\beta\sqrt{\nu + 1} \quad (1.9)$$

where ν is the vibrational quantum number and $b = \hbar/4\pi c\tilde{\nu}$ is the quantum-mechanical analog of the amplitude of a classical oscillator. From Eq. (1.8), it is seen that at least one component of the polarizability tensor should be non-zero for the transition to be Raman active. The intensity ratio of Stokes to anti-Stokes lines is,

$$\frac{I_{Stokes}}{I_{anti-Stokes}} = \left(\frac{\tilde{\nu}_0 - \tilde{\nu}_k}{\tilde{\nu}_0 + \tilde{\nu}_k} \right) \exp\left(\frac{hc\tilde{\nu}}{kT}\right) \quad (1.10)$$

The Raman scattering cross-section representing the cross-sectional area of the molecule interacting with the incident electromagnetic radiation is given by,

$$\sigma(i \rightarrow f) = \frac{8\pi\omega_s^4}{9\hbar c^4} \left| \sum_j \left(\frac{\langle \alpha_{ij} \rangle \hat{e}_L \langle \alpha_{jf} \rangle \hat{e}_s}{\omega_{ij} - \omega_L - i\gamma_L} + \frac{\langle \alpha_{ji} \rangle \hat{e}_L \langle \alpha_{jf} \rangle \hat{e}_s}{\omega_{jf} - \omega_L - i\gamma_j} \right) \right| \quad (1.11)$$

where \hat{e}_L and \hat{e}_s are unit vectors, which represent the polarisation of the incident laser beam and the scattered light. The sum extends over all molecular levels j with the homogenous width γ . The initial and final are connected by two-photon transitions. The probability term in the Raman scattering cross-section should be non-zero for any vibration to be Raman active.

1.2. Information from a Raman spectrum

The characteristic Raman frequencies represented by a band position at a specific Raman shift originates from the specific functional groups and chemical species (Fig. 1.4). The whole Raman spectrum as a set of these multiple Raman bands is a molecular fingerprint indicating the composition of the material. The intensity of the Raman band is directly proportional to the concentration of the corresponding chemical species. Also, it is affected by its scattering cross-section of the various vibrational modes, which varies amongst different chemical species. The FWHM of the Raman band indicates the quality of the material, i.e. the crystallinity, the defects, doping and structural disorders. It also indicates the lifetime of the phonon or vibron

in the system. The shift in the position of a Raman band is directly related to external chemical and physical factors like temperature, pH, pressure, stress, strain, or deformation changes. Overall, comparing two Raman spectra under different conditions can give us a picture of the effects on a molecule. The polarisation dependence of a Raman peak provides insight into the crystal symmetry and orientation of the crystals in the sample.

Biomaterials have a large number of atoms, indicating many active vibrational modes. These modes result in a complex spectrum, where a peak's origin cannot be directly correlated to a single vibrational mode but rather a result of many similar vibrations throughout the molecule. When affected by environmental parametric changes, these vibrations are reflected as changes in the Raman spectral features.

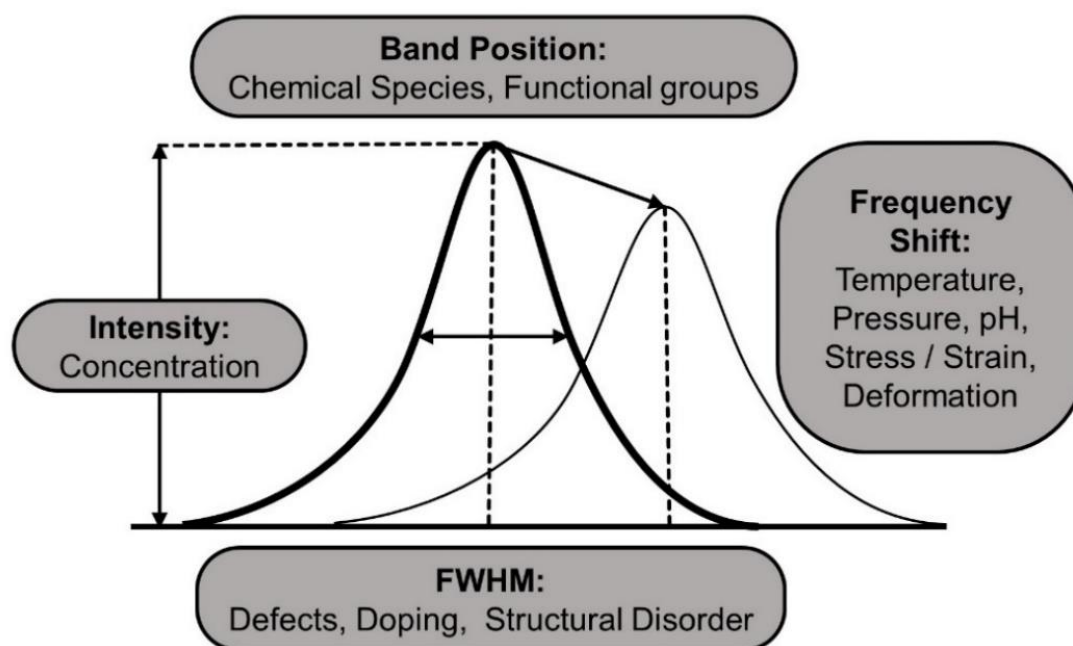


Fig. 1.4. Various components of a Raman spectrum provide insight into different aspects of the effect on the molecular vibrational modes.

Though the Raman spectrum is very informative, interpretation of the data gets difficult if there is no sufficient spectral signal. This happens because Raman scattering is a weak phenomenon. When light scatters off from a sample, most of the light is elastically scattered, making the majority of the photons Rayleigh scattered. A tiny fraction, i.e., one in 10^7 photons, is only Raman scattered, thus making it a weak phenomenon.

At lower concentrations of the sample, normal laser Raman scattering suffers from low scattering cross-section of the molecules. The fluorescence cross-section of a molecule is typically of the order of 10^{-17} cm² per molecule, and the non-resonant Raman cross-section is around 10^{-26} cm² per molecule.⁹ Thus, Raman signals are sometimes masked by the fluorescence background for fluorescent molecules. Many techniques can be employed to improve the Raman signal. Some of them include a) changing the wavelength of the laser to avoid the fluorescence background, b) Use a choice of lasers to excite specific vibrational modes of the molecule, thus producing resonant Raman spectra and c) using surface-enhanced Raman spectroscopy (SERS) to improve the intensity by a many-fold.

1.3. Surface-enhanced Raman spectroscopy

In 1974, Fleischmann et al. observed an increase in the Raman intensity of molecules on roughened electrodes, thus establishing the Surface-enhanced Raman scattering phenomenon.¹⁰ In the following years, other groups also observed and interpreted these observations.^{11, 12} From their research it was evident that, plasmonic nanostructures of noble metals like gold, silver and copper show higher SERS enhancement than metals like aluminium, lithium, and sodium.¹³ Colloidal metallic nanoparticles, arrays of metallic nanoparticles on a surface and metal electrodes have proven to be good SERS substrates.

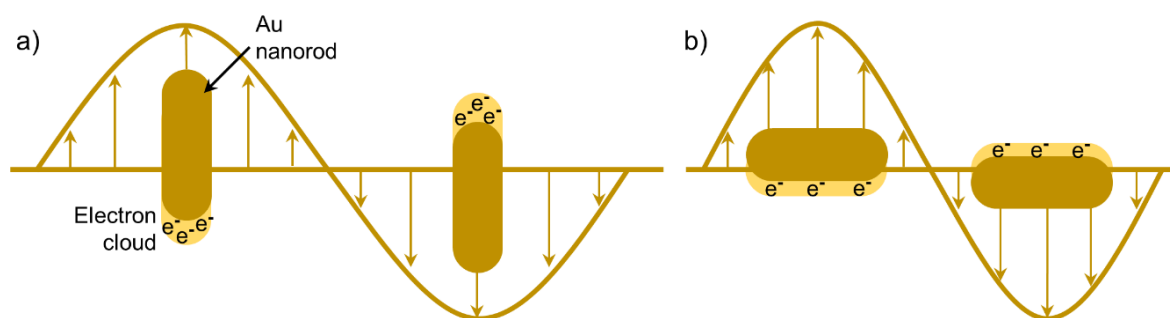


Fig. 1.5. Surface plasmon polaritons in metallic gold nanorods, a) Longitudinal charge density waves and b) Transverse charge density waves.

SERS is a technique of enhancing Raman scattering cross-section by adsorption of molecules on metallic nano surfaces.¹⁴ The collective oscillations of electrons on the surface of these nanostructures are called surface plasmons. The incident

electromagnetic radiation excites the surface plasmons, resulting in surface plasmon polaritons (Fig. 1.5). When the plasmons come into resonance with the frequency of the incoming radiation, there is surface plasmon resonance. Their nano surfaces exhibit atomic-scale roughness and act as plasmonic surfaces to the molecules in their close vicinity. If the dimensions of these plasmonic nanostructures are lesser than the wavelength of light, localized surface plasmon resonance is created. The excited surface plasmon creates an oscillating dipole that produces an intense electromagnetic field close to the surface. A molecule in the proximity of the nanostructure experiences this highly intense field, leading to SERS. The scattering cross-section is now comparable to fluorescence, thus enhancing the Raman intensities.¹⁴

1.3.1. Factors affecting SERS enhancement

In standard Raman spectroscopy, the average Raman intensity of a molecule is directly proportional to the laser power density and the Raman scattering cross-section. However there are many factors that affect the SERS intensity. The ones that directly affect the Raman spectrum include the strength and wavelength of the laser, collection optics and scattering geometry, and intrinsic Raman cross-sections of the molecule. The parameters specific to SERS include the analyte adsorption properties and the SERS substrate's material, geometry, dimensionality, and orientation with respect to incident light. An understanding of the mechanism of how the enhancement happens is essential. This allows one to tune these factors and thus produce high SERS enhancements. Traditionally, SERS enhancement can be explained using two mechanisms, electromagnetic and chemical enhancements.¹⁵ The first one involves enhancement from the local electric field, while the latter is a result of enhancement in polarizability of the molecule due to chemical effects such as charge transfer excited states and the formation of resonant intermediates.

1.3.1.1. Electromagnetic enhancement

Electromagnetic enhancement occurs when the molecule is in close vicinity of the plasmonic nanoparticle. When an appropriate incident photon of energy $h\nu$

incident interacts with the nanoparticle, the plasmon gets excited. On resonance, the electron oscillations in the nanoparticle induce a strong electric field on its surface, due to the localised surface plasmon resonance (LSPR). This strong electric field transfers energy to the molecule, thus exciting it to a higher virtual energy state. As it returns to the ground state, it transfers the energy back to the nanoparticle, which gets released as $h\nu$ -scattered. This photon can get scattered or be reabsorbed and further enhanced. A schematic describing the various paths of the energy beginning from the absorption of the photon to the scattering of another photon is represented in Fig. 1.6.

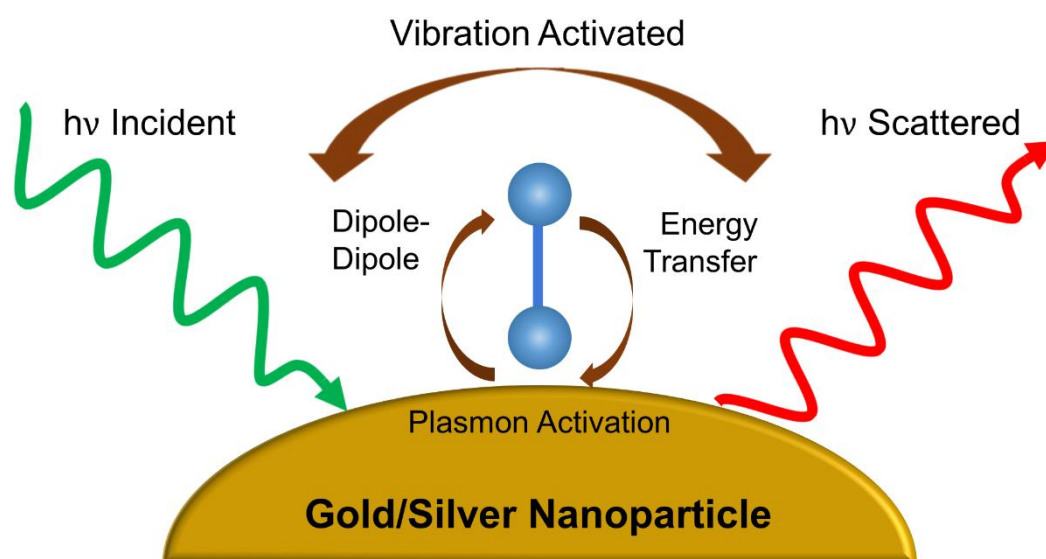


Fig. 1.6. Schematic representation of the Electromagnetic enhancement mechanism in SERS.

According to the Drude model, metals are considered as periodically distributed static positive charges that are surrounded by a sea of electrons. When the electrons are displaced, a momentary dipole can be induced. This dipole oscillates with its own characteristic frequency and generates an electric field (Fig. 1.7). When a molecule is placed in its vicinity, it can experience the enhanced electric field. To understand the enhancement, we can consider a metallic sphere that is excited by an oscillating electric field and observe the energy of the scattered radiation.

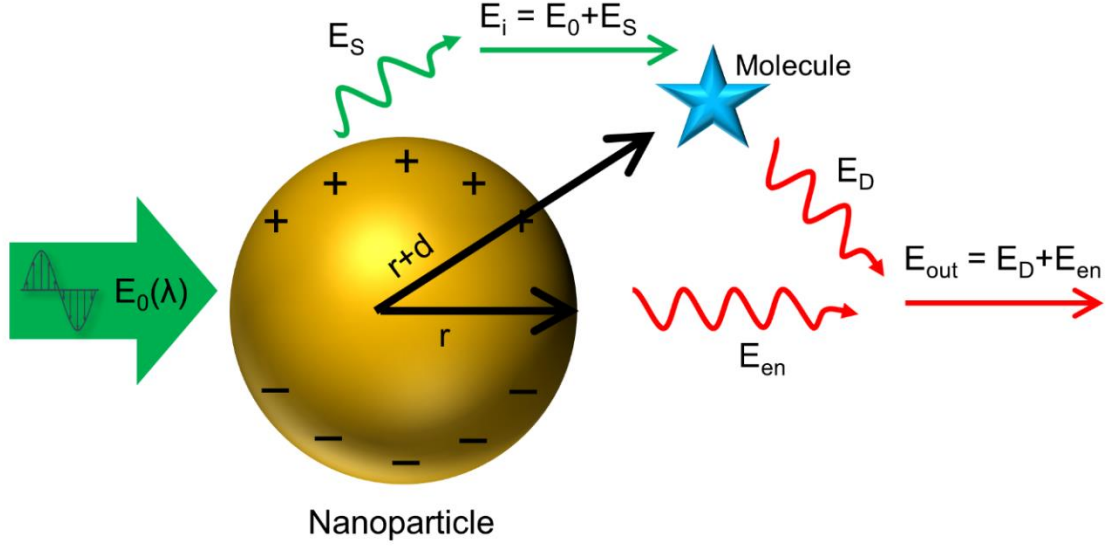


Fig. 1.7. An incident time-varying electric field E_0 on a metallic nanoparticle results in an electromagnetic enhancement. (Inspired from Siddhanta, S. et al,¹⁶)

When an electric field E_0 is incident on a metallic nanoparticle, the resultant field that is incident on the molecule is $E_i = E_0 + E_S$, where E_S is due to the surface plasmon oscillations. E_S depends on the sphere's radius r and the distance of the molecule from the sphere, d . It also depends on the dielectric constant ε of the metal and the incident light's electric field, E_0 . From Drude's model, E_S is given by,

$$E_S = \frac{\varepsilon - \varepsilon_0}{\varepsilon + 2\varepsilon_0} \left(\frac{r}{r+d} \right)^3 E_0 \quad (1.12)$$

The resultant field, which is now incident on the molecule, E_i induces a dipole μ_i in the molecule. This resultant dipole is directly proportional to the electrical field E_i and the polarizability tensor of the molecule α . As discussed earlier, the molecular vibrations combine with the oscillating field of the incident E_i , resulting in three kinds of frequencies Rayleigh, Stokes and anti-Stokes scattering. In the case of Stokes scattering, the radiation from the dipole, E_D , is redshifted. The resultant outgoing field is given by $E_{out} = E_D + E_{en}$, where E_{en} is the additional component of enhancement due to the elastic scattering producing further induced dipole in the molecule. The enhancement in the field is thus given by the ratio of the field amplitudes, E_i/E_0 and

E_{out}/E_0 for the laser and the Raman scattered field, respectively. The SERS enhancement is the product of the laser and the Raman scattered fields,

$$G_{SERS} = |A(\nu_L)|^2 |A(\nu_S)|^2 = \left| \frac{\varepsilon(\nu_L) - \varepsilon_0}{\varepsilon(\nu_L) + 2\varepsilon_0} \right|^2 \left| \frac{\varepsilon(\nu_S) - \varepsilon_0}{\varepsilon(\nu_S) + 2\varepsilon_0} \right|^2 \left(\frac{r}{r+d} \right)^{12} \quad (1.13)$$

G_{SERS} is proportional to the fourth power of the local electric field at the vicinity of the metallic nanostructure and is strong when the scattered and plasmon fields are in resonance. G_{SERS} is maximum when the real part of $\varepsilon(\nu)$ is equal to $-2\varepsilon_0$ and the imaginary part is negligible. Metals like silver (Ag), gold (Au) and copper (Cu) satisfy the conditions at visible wavelengths where most molecules' vibrational modes are activated.¹⁷ Thus, they show high SERS enhancements in visible light. The enhancement factor is distance-dependent, and it decays as $\left(\frac{r}{r+d}\right)^{12}$. So, the closer the molecule is to the metallic nanoparticle, the higher is the enhancement factor. For electromagnetic enhancement to occur, the molecule's presence in close proximity is essential, while it is not a compulsion for the molecule to form a chemical bond with the metallic nanoparticle. Hence, electromagnetic enhancement is an analyte independent phenomenon. It has been shown that the enhancements of the order of 10^6 - 10^9 have been obtained from its contribution.

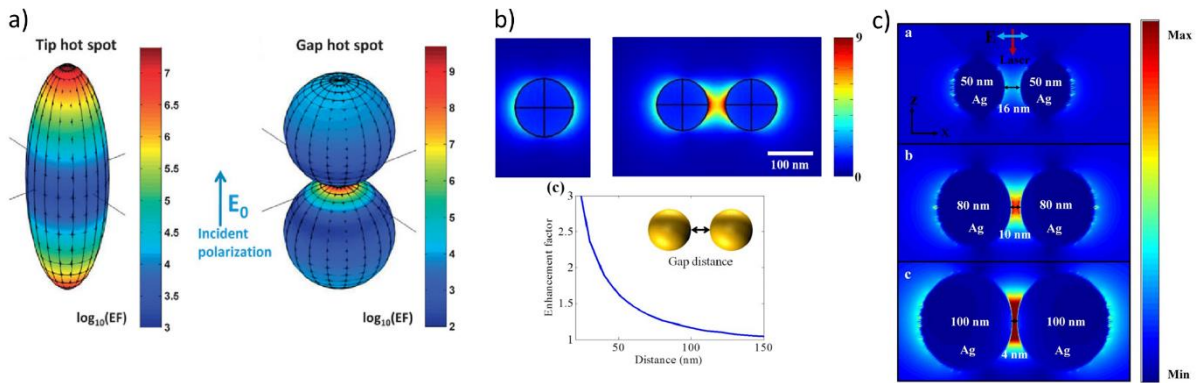


Fig. 1.8. a) Enhanced electric fields in the tips and the gaps of nanoparticles, b) comparison of the field in single nanoparticle vs gap in two, c) Distance dependent electric field enhancement. (Figures adapted from¹⁸⁻²⁰) Copyright © 2013, The Materials Research Society; Copyright © 2019, Xin Zhang et al.

The enhancement factor also depends on the specific geometries of the nanostructures or more localized areas with extremely high enhancements. Such localised areas with high SERS enhancements are called “hotspots”. These hotspots are formed in the gaps between closely spaced nanostructures and at sharp tips. (Fig. 1.8a).¹⁹ It has been calculated that the electric field strength in this gap between adjacent particles is a few orders higher in magnitude than on the surface of a single isolated plasmonic nanoparticle, thus resulting in much larger SERS enhancement factors (Fig. 1.8b).¹⁸ As seen in Fig. 1.8c, the closer the nanoparticles are, the larger the enhancement in the electric field created in the gap between them.²⁰ In this thesis, AgNPs have been seen to form aggregates in the presence of spermine tetrahydrochloride or by placing them on ice, thus increasing the number of formed hotspots.

1.3.1.2. Chemical Enhancement

Chemical enhancement occurs when a bond is formed between the molecule and the surface of the metal nanoparticle. It is also known as a first layer effect and forms an adsorbate-surface complex. Electronic coupling can occur when a molecule is in contact with the surface of the metal. This causes an enhancement in the Raman cross-section of the adsorbed molecule as compared to its free form in the case of normal Raman. This charge transfer leads to the broadening and shifting of the electronic level in the molecule, similar to the resonance Raman effect.

The energies of the highest occupied molecular orbital (HOMO) and the lowest unoccupied orbital (LUMO) are symmetric to the Fermi level of metal (Fig. 1.9). The shift in Fermi Level changes the resonance conditions. When a photon is absorbed by the metal, it results in a hot electron state. This hot electron transitions from the HOMO to the LUMO of the molecule. When the metal returns to its ground HOMO state, it does so by emitting a Stokes photon.

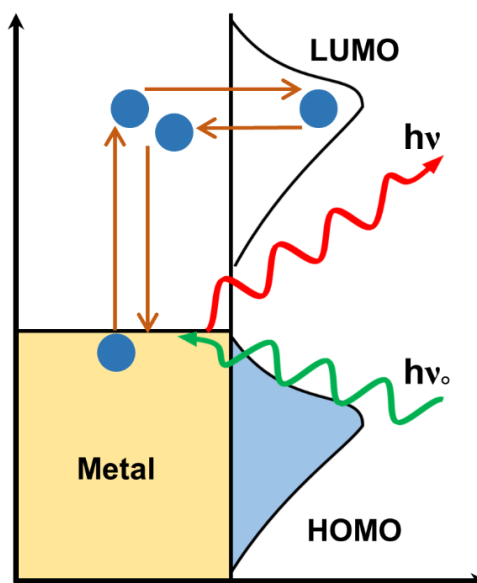


Fig. 1.9. Schematic representation of the four-step process of the photon-driven charge transfer model for a molecule adsorbed on a metal surface resulting in chemical enhancement.

There are mainly three kinds of contributions for chemical enhancement contribution to the SERS enhancement factor. They include, (i) resonance Raman (RR) effect, which occurs when the energy of the incident light matches an electronic transition in the molecule (10^3 – 10^6 contribution), (ii) a charge-transfer (CT) effect where the incident light resonates with a metal-molecule or molecule-metal transition (10 – 10^4 contribution), and (iii) a non-resonant chemical effect that occurs due to ground-state orbital overlap between the molecule and the metal (≤ 10 – 100 contribution).²¹ The charge transfer mechanism is greatly dependent on the geometry, bonding, and the energy levels of the molecule, though it is very short-ranged (0.1–0.5 nm).

1.3.2. Nanostructures for SERS

For large enhancement factors, an intense surface plasmon resonance and nano-scale roughness of the metallic surface play a significant role. The shape, component, size, and interparticle spacing of the metallic nanostructures are critical factors influencing the SPR and electromagnetic fields. As for metallic nanoparticles, the surface to volume ratio is high, indicating that most of their atoms are on the surface. This feature provides both surface roughness and increases the chances of surface plasmon resonance. As seen earlier, the most commonly used metals for preparing

nanostructures for SERS include Ag and Au, and sometimes Cu, Pd and Pt. They have their surface plasmon resonances in the visible region.

Of the many processes that can be used to develop SERS substrates, the solution-based synthesis is the most common. This widespread use is owed to the simplicity and ease of their preparation. This method can result in dispersed particles and aggregated or assembled systems. Various nanostructures can be made from simple nanospheres to more complex nano-stars,²² nanorods,^{23, 24} nano-urchins,^{25, 26} and core-shell nanoparticles.^{27, 28} In SERS literature, there are two extensively used nanoparticle synthesis techniques. The most popular method is the citrate-reduction method by Lee-Meisel.²⁹ Here, silver nitrate is reduced by sodium citrate, where the citrate ions act as reducing agents and as capping agents to provide stability to the nanoparticles.²⁹ The other method is by Creighton et al., where sodium borohydride is used as the reducing agent.³⁰

There are many other advanced techniques that result in more intricate nanostructures. Some solid-supported SERS substrates include nanotriangle arrays,^{31, 32} metal film over a nanosphere substrate,³³ and metal nanoparticles over Si nanopillars.³⁴ These have advantages for single-molecule detection ultralow volumes and can be integrated with microfluidic setups. They suffer from higher costs for using nanolithography techniques. The imaging capabilities of a scanning probe microscope can be combined with the high enhancement properties of SERS to give a new technique called tip-enhanced Raman spectroscopy. Here, either a metallic film,³⁵ nanospheres³⁶ and nanocubes³⁷ are used to modify a probe to obtain highly sensitive and spatially resolved Raman information from a sample. Spatially organised nano-architectures obtained through controlled aggregation by forming nanometre gaps between different components are an essential feature for single molecular detection SERS substrates.

Many techniques allow for the construction of different nanostructures and their assemblies for SERS applications. Depending on the final aim of the SERS study, either

simpler solution-based nanoparticles or sophisticated nano-architectures can be used. In this thesis, simple citrate reduced nanoparticles based on the Lee-Meisel method have been used for all the SERS studies. These nanoparticles have been shown to provide enhancement factors of the order of 10^4 to 10^6 .³⁸

1.4. Raman spectroscopy for biosystem analysis and detection

Biomolecules are essential building blocks of life. Studying living creatures' cellular activities and how these processes relate to the organism's functioning is a key focus of biochemistry. Over the last century, biochemistry has allowed us to understand atoms and biomolecules that make up living organisms, the central dogma around which biological information is transferred. This allows for a better understanding of ourselves and our surroundings. The functional homeostasis of a cell is maintained by various biomolecules, including nucleic acids, proteins, lipids, and carbohydrates. Impressively, each of these biomolecules has a distinctive chemical identity and, as a result, a specific molecular structure. Any biological system or material is a combination of these various biomolecules; hence it becomes essential to understand their structure and properties.

Studying the function of membrane protein drug targets for certain diseases, we can establish the link between structural biology and medical applications.³⁹ Structural studies at the molecular level provide an in-depth understanding of these biomolecules and can reflect in establishing new medical applications or strategies. The positions, widths, and heights of the Raman bands are sensitive to the molecular structure that can be affected by minor perturbations induced by intramolecular and intermolecular interactions. Hence, Raman spectroscopy has been proved competent in establishing the connection between the structure and function of biomolecules. The overall Raman spectrum can provide an idea of the chemical composition of the biosystem under study, thus making it useful for characterisation and detection.

Raman, NIR and fluorescence spectroscopy are the most important for biomedical applications of the many optical molecular spectroscopy techniques.

Fluorescence spectroscopy depends on extrinsically applied labels to provide specific information. Though sensitive, it is disadvantageous as it can't be used to explore structural changes of biomolecules or study *in situ* systems due to the limited number of non-toxic dyes. NIR spectroscopy has high permeability, thus allowing for studying whole cells or bodies *in situ* and therefore examining internal structures non-destructively. It offers high sensitivity and specificity, but is disadvantageous due to highly intense water bands in an NIR spectrum of a sample containing water, which most biomolecules do.

With advancements in lasers, optical instrumentation and a better understanding of the information obtained by a Raman spectrum, Raman spectroscopy has become an essential tool for chemists, physicists, and biologists over the past many years. The most significant benefit of this technique lies in its high sensitivity to subtle molecular changes that can occur due to any external parameters. As this technique is simple to use and doesn't require any sample preparation, it is versatile of the range of samples, this can be used for. Additionally simple spectrometers can be built for its study and the sample can be studied *in situ* under a variety of extreme conditions such as temperature, pressure, and pH. Further, water having very weak Raman modes allows for studying biosystems in their natural environment, without solvent interference. And as the technique is non-destructive, it can also be used for many *in situ* measurements. Further, the spatial resolution of micro Raman spectroscopy in the low micrometre scales and its ability to probe samples under *in vivo* conditions allows new insights into systems like living single cells.⁴⁰ Over the past many years, research involving isotope labelling, model compounds, site-directed mutagenesis and normal coordinate analysis has led to the identification of so-called marker bands.⁴¹ The characteristic vibrational bands and these marker bands act as bridges in establishing the correlations between the spectrum and the composition and structures.

Most biosystems are comprised of proteins, lipids, nucleic acids, and carbohydrates. Hence, a Raman spectrum of a biosystem results from vibrations from

all these components combined into one. Broadly, the biochemical information in a spectrum can represent the classes of chemical bonds but can get tricky due to its complex nature. Analysing these complex spectra and assigning the various peaks can be challenging but can yield precious information regarding the structure and dynamics of the system under study. As it gets challenging to provide spectral assignments to specific bonding schemes, concentrations or geometries, biological Raman spectra are analysed as phenomenological signatures. Irrespective of the complexities, some Raman bands can be assigned to a direct chemical bond, for example, the in-phase ring breathing of Tryptophan around 750 cm^{-1} and C-H stretching around 2900 cm^{-1} . The differences amongst the biological Raman spectra are mainly due to the subtle differences that occur due to compositional and structural variations. While understanding and assigning the peaks is beneficial, it is not required to chemically evaluate the entire spectrum to utilise it for comparison under different conditions. It is essential to understand and assess phenotypic variability resulting from external environmental factors like temperature, age, nutrients and intensive properties like structural and organism composition, which can reflect as small spectral differences.⁴²

The Raman spectra of biological samples can be divided into three spectral regions. The spectral range 500 to 1800 cm^{-1} is the “fingerprint region” consisting of sharp, localized spectral features unique to the constituent molecules.⁴³ These allow for classifying samples and chemometric analysis. The bands in this region originate from vibrations of slightly larger atoms like carbon, oxygen, nitrogen, sulphur, or complexes of many hydrogen atoms. Independently vibrating hydrogen atoms vibrate with much higher energies than other bonds, as the smaller mass of hydrogen causes an overall smaller reduced mass for the system. These appear in the 2500 to 3600 cm^{-1} spectral range and are called the “high wavenumber region”.⁴³ In between these two regions lies the “silent region”, ranging from 1800 to 2500 cm^{-1} . This region is mostly empty and has no biomolecular contributions, except for a few exceptions.⁴³ In this thesis, the fingerprint regions of the biosystems under study have been used to characterise them as well as monitor changes due to external environmental parameters.

As seen earlier, spontaneous or standard Raman spectroscopy produces very low signals. If the sample doesn't have a large Raman scattering cross-section like most biomolecules, even longer acquisitions result in low signals. Over the recent past, many new variations of the Raman technique have been developed to overcome this limitation. Exploiting the use of metallic nanostructures for their localised surface plasmons allows for a many-fold increase in the Raman spectral intensities. Techniques like SERS (surface-enhanced Raman spectroscopy) employs silver or gold nanostructures that significantly increase the Raman intensity of the molecules in their close vicinity. This can allow single molecular detection^{44, 45} and is useful for trace material analysis.⁴⁶ SERS has an added benefit of quenching the autofluorescence generated by biomolecules, thus reducing the background obtained in the spectrum.⁴⁷ Within biosciences, SERS has been used to study from single biomolecules to cells to tissues. They have applications in cell studies, medical applications and toxicology. Various cells and organelles like lymphocytes, haemoglobin, red blood cells, mitochondria, cardiomyocytes, neuronal cells, etc., have been detected using SERS.⁴³ Additionally, SERS has been used to detect other biosystems like bacteria, proteins, and antigens.

Resonant Raman spectroscopy occurs when the laser wavelength is chosen to match the energy of the electronic transition, thus causing resonance. In the presence of plasmonic nanostructures and the optimal choice of laser, surface-enhanced resonance Raman spectroscopy (SERRS) can be achieved, producing an even higher enhancement in the Raman signal. When Raman is combined with an optical microscope, molecular properties of the sample can be extracted with diffraction-limited spatial resolution. Raman imaging can be performed by spatially mapping the sample and collecting Raman spectra at different points. When a Raman spectrometer is integrated with a scanning probe microscope, the optical coupling brings the excitation laser to the functionalized probe (metallic tip), thus providing Raman spectra with local enhancement and an increased spatial resolution. These different techniques can be employed to improve the signal of the biosystems under study.

Specific techniques can be chosen depending on what aspects need to be studied. There are many other variations of Raman spectroscopy that can be explored, including spatially offset Raman spectroscopy (SORS), transmission Raman spectroscopy (TRS), stimulated Raman spectroscopy (SRS), coherent anti-stokes Raman spectroscopy (CARS), and selective scanning Raman spectroscopy (SSRS).^{43, 48} This thesis mainly uses SERS for detection of extracellular vesicles, and mutations in the DNA of the tomato plants. We have used drop-coated deposition Raman spectroscopy (DCDR), which allows for a higher concentration of the protein to be present under the laser spot for the lysozyme aggregation studies.

A common issue in biological Raman spectroscopy is the presence of unnecessary background. When an external light source illuminates a biomolecule or tissue, it can generate a broadband natural emission called autofluorescence in the visible range. It is one of the major noise sources in Raman spectroscopy of biological tissues. Depending on the varying constituents of the system, its strength and spectral shape can vary. This can be avoided using different laser wavelengths. An optimal choice of the laser wavelength and power of the laser source has to be made to prevent autofluorescence while simultaneously obtaining good Raman spectra. The other sources of noise can be from the instrument or external sources. Noise like dark current and thermal noise are generally eliminated using thermoelectrically cooled CCD detectors. Shot noise, originating from the particle nature of light, contributes to $n^{1/2}$ for a measurement of n counts. So, the signal to noise (S/N) ratio can be improved by increasing exposure time and averaging multiple cycles. The intrinsic fluorescence emission is several orders of magnitude higher than the Raman signal for biosystems. Most commonly, a polynomial fit is used to remove the background effects as seen in Fig. 1.10.

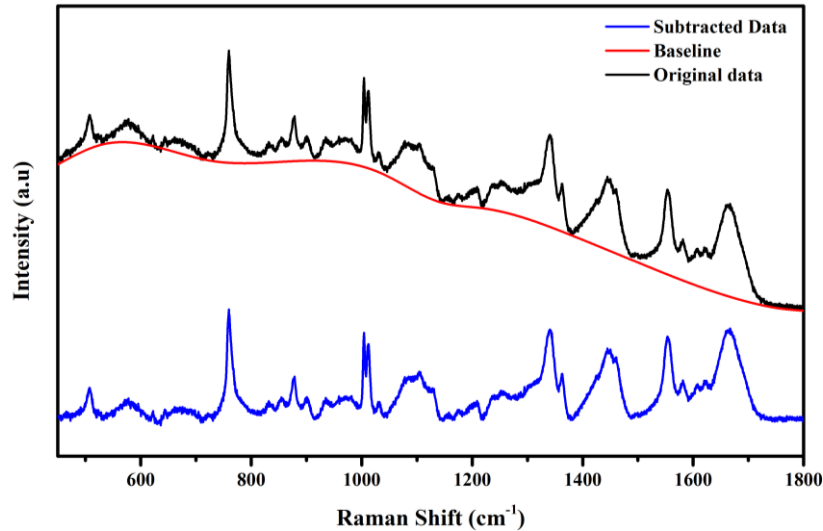


Fig. 1.10. Baseline correction of the background from the original data.

When high energy cosmic rays impact a charge-coupled device detector, sudden spurges of energy can appear as very narrow spikes on the Raman spectra.⁴⁹ When the Raman intensities are low, and we collect spectra for long durations, there are higher chances of cosmic ray interactions, leading to a larger number of cosmic spikes that appear on the spectra. Most spectral collection software have a tool to remove the cosmic radiation before further data processing.

Raman spectra from the same sample could have different intensities when experimental parameters vary. Normalized spectra allow for easier comparison of the intensities of peaks on different Raman spectra. These are achieved by peak normalisation or area normalization. In the peak normalization, the intensity of the spectra is divided by the intensity corresponding to the central frequency of a reference Raman band.⁵⁰ This method assumes that the reference band does not change amongst the spectra and is not affected by any changes in the sample or external conditions. Area normalisation is the best process when the spectra do not share a common or unaffected band. The intensity at each frequency is divided by the square root of the sum of squares of all intensities. Here, the total area under the curve becomes 1. ⁵⁰ Most Raman spectra are fitted using Gaussian, Lorentzian or a combination like the Voigt curve for obtaining various spectral parameters like peak position, FWHM and intensity. Sometimes a spectral peak originates from multiple contributions (like the

Amide I band of the proteins), and spectral deconvolution is required to assess the various components. So along with improved experimental conditions like higher and multiple acquisitions, basic data processing, including background correction and spectra normalisation, allows for better analysis of the biological Raman spectra.⁵⁰

The most fundamental way of analysing every individual feature in the Raman spectrum retains the most chemical information. Sometimes minor errors in background corrections can affect the absolute intensity analysis of the Raman band. In such cases, the ratios of Raman bands can be used as indicators. Though Raman spectroscopy is very accurate, the spectral differences between samples can vary only slightly. Chemometrics is a chemical discipline that uses mathematical, statistical, and other formal logic-based methodologies to develop or choose optimal measuring methods and experiments and analyse chemical data to offer the most relevant chemical information. Hence this discipline helps categorise and differentiate systems based on subtle changes in the Raman spectra.^{42, 43, 48, 50} These methods are diverse and use different approaches to extract specific information from the data. There are two kinds of such techniques, multivariate classification and multivariate regression. The unsupervised pattern recognition techniques, principal component analysis (PCA) and cluster analysis (CA) do not require prior knowledge about the samples. They are used to look at the differences and similarities between spectra. Other techniques like linear discrimination analysis (LDA) and artificial neural networks (ANN) are supervised techniques that require prior knowledge of the samples. Multivariate regression techniques can be used, which are generally applied to analyse one or multiple molecules of a complex sample that possesses overlapping spectroscopic signals. These include principal component regression (PCR), logistic regression (LR), partial least squares regression (PLS). In this thesis, we have used PCA for categorising the Raman spectra corresponding to extracellular vesicles extracted from various cell lines. SERS combined with PCA has been used for detecting lung cancer, gastrointestinal cancer, oral cancer, skin cancer, and various skin diseases.⁵⁰

Raman spectroscopy has taken a long course and has developed importance to be a mainstream tool. Over the past years, this has been achieved by improving the instrumentation, allowing for obtaining a more resolved and better quality spectrum. The transition from mercury arc lamps to lasers that provide stable and intense beams of radiation was a big turning point. A combination of notch filter and high-quality grating monochromators in dispersive instruments provide a better Raman spectrum. Based on the area of use, there are two kinds of instruments: lab-based spectrophotometers and in-field or *in situ* spectrophotometers.⁵¹ They both work on the same principle, and their differences come from the instrument's versatility and size and relative cost of the components. *In situ* Raman analysis is possible nowadays with the development of miniature tabletop or portable Raman spectrometers. Remote sensing has been possible by integrating the spectrometer with optical fibres.⁵² In this thesis, we have briefly explored the construction of a miniature Raman spectrometer that can find on-site applications.

Our research group over the past years have established themselves in exploiting Raman spectroscopy for biosystem analysis and detection. Research has been carried out on various fronts, including understanding protein-protein, protein-drug interactions and using SERS to detect biomolecules. The ability of SERS to differentiate between structurally homologous proteins Aurora A and Aurora B validates the strength of Raman spectroscopy.⁵³ Molecular dynamics assisted SERS was utilized to probe the effect of Mg^{2+} and Ca^{2+} ions on the structure of Kpn1 protein.⁵⁴ Vibrational modes associated with aromatic amino acids near metal-binding sites, dimer assembly, partial loss of secondary structure were probed. 92% structurally similar kinases, JNK1 and JNK3, oriented themselves differently on the surface of a nanoparticle, and thus through SERS, the ligand-binding site was established.⁵⁵ From such studies, the field of drug discovery can benefit by screening the binding of small molecules on proteins. With the help of resonance Raman spectroscopy, the structure of BLA (biliverdin IXa) in native and wildtype Sandercyanin protein was characterised.⁵⁶ In the absence of the crystal structure, the SERS studies of p300, a

large multidomain transcriptional coactivator protein, gave structural insights and the effect of trifluoromethyl phenyl benzamides on the histone acetyltransferase (HAT) activity of p300.^{57, 58} Not just the various aspects of protein-protein, protein-drug interactions, but research in the direction of biomolecular detection was also explored. A new strategy to detect HIV-1 subtype using SERRS was developed, tested and patented.⁵⁹

1.5. Scope of this thesis

As seen in the last section, Raman spectroscopy and its variations have a wide range of biological applications. It is evident that Raman spectroscopy is not just in its scientific research stages but is progressing towards becoming a technique that can be used for biomedical applications. Here in this thesis, four work chapters (chapters 3-6) cover various aspects of the same. Chapter 3 uses Raman spectroscopy to understand a molecular picture of the aggregation process of the hen egg-white lysozyme and human lysozyme in an alkaline environment. Here, DCDR spectroscopy is employed to concentrate the protein, to get an improved Raman signal. Chapter 4 covers the use of SERS and a multivariate analysis tool, PCA, to differentiate extracellular vesicles extracted from various cell lines under autophagic and other conditions. Along with this, some biochemical insights were obtained by analysing the Raman spectra. Chapter 5 discusses the use of SERS to detect a Raman active molecule that is a part of a probe that has been designed for detecting SNPs and indels in tomato plants. This has been further extended to multiplexing studies that can be employed in detecting a random sample of DNA that has been extracted. Finally, in chapter 6, a miniature Raman spectrometer was constructed to allow for on-site studies of various applications. The thesis ends with an outlook that suggests various other aspects that can be developed in the future, based on the present studies of this thesis.

1.6. References

1. Tissue, B.M. Ultraviolet and Visible Absorption Spectroscopy. *Characterization of Materials*, 1 (2012).
2. Rayleigh, L. XXXIV. On the transmission of light through an atmosphere containing small particles in suspension, and on the origin of the blue of the sky. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 47, 375 (1899).
3. Smekal, A. Zur Quantentheorie der Dispersion. *Naturwissenschaften* 11, 873 (1923).
4. Raman, C.V. A Change of Wave-length in Light Scattering. *Nature* 121, 619 (1928).
5. Raman, C.V. & Krishnan, K.S. A New Type of Secondary Radiation. *Nature* 121, 501 (1928).
6. Placzek, G. Rayleigh-Streuung und Raman-Effekt (Akademische Verlagsgesellschaft, Leipzig, 1934).
7. Kramers, H.A. & Heisenberg, W. Über die Streuung von Strahlung durch Atome. *Zeitschrift für Physik* 31, 681 (1925).
8. Dirac, P.A.M. & Bohr, N.H.D. The quantum theory of the emission and absorption of radiation. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* 114, 243 (1927).
9. Kneipp, K. et al. Single Molecule Detection Using Surface-Enhanced Raman Scattering (SERS). *Physical Review Letters* 78, 1667 (1997).
10. Fleischmann, M., Hendra, P.J. & McQuillan, A.J. Raman spectra of pyridine adsorbed at a silver electrode. *Chemical Physics Letters* 26, 163 (1974).
11. Albrecht, M.G. & Creighton, J.A. Anomalously intense Raman spectra of pyridine at a silver electrode. *Journal of the American Chemical Society* 99, 5215 (1977).
12. Jeanmaire, D.L. & Van Duyne, R.P. Surface raman spectroelectrochemistry: Part I. Heterocyclic, aromatic, and aliphatic amines adsorbed on the anodized silver electrode. *Journal of Electroanalytical Chemistry and Interfacial Electrochemistry* 84, 1 (1977).
13. Zeman, E.J. & Schatz, G.C. An accurate electromagnetic theory study of surface enhancement factors for silver, gold, copper, lithium, sodium, aluminum,

- gallium, indium, zinc, and cadmium. *The Journal of Physical Chemistry* 91, 634 (1987).
14. Nie, S. & Emory Steven, R. Probing Single Molecules and Single Nanoparticles by Surface-Enhanced Raman Scattering. *Science* 275, 1102 (1997).
 15. Campion, A. & Kambhampati, P. Surface-enhanced Raman scattering. *Chemical Society Reviews* 27, 241 (1998).
 16. Siddhanta, S. & Narayana, C. Surface Enhanced Raman Spectroscopy of Proteins: Implications for Drug Designing. *Nanomaterials and Nanotechnology* 2, 1 (2012).
 17. Baibarac, M. et al. SERS spectra of polyaniline thin films deposited on rough Ag, Au and Cu. Polymer film thickness and roughness parameter dependence of SERS spectra. *Synthetic Metals* 96, 63 (1998).
 18. Chung, T., Lee, S.-Y., Song, E.Y., Chun, H. & Lee, B. Plasmonic Nanostructures for Nano-Scale Bio-Sensing. *Sensors* 11 (2011).
 19. Le Ru, E.C. & Etchegoin, P.G. Quantifying SERS enhancements. *MRS Bulletin* 38, 631 (2013).
 20. Zhang, X. et al. Organic Molecule Detection Based on SERS in Microfluidics. *Scientific Reports* 9, 17634 (2019).
 21. Kumar, S., Kumar, P., Das, A., & Pathak, C. S. in Recent Advances in Nanophotonics - Fundamentals and Applications (ed. In M. Kahrizi, P.A.S.) (IntechOpen, 2020).
 22. Garcia-Leis, A., Garcia-Ramos, J.V. & Sanchez-Cortes, S. Silver Nanostars with High SERS Performance. *The Journal of Physical Chemistry C* 117, 7791 (2013).
 23. Saute, B., Premasiri, R., Ziegler, L. & Narayanan, R. Gold nanorods as surface enhanced Raman spectroscopy substrates for sensitive and selective detection of ultra-low levels of dithiocarbamate pesticides. *Analyst* 137, 5082 (2012).
 24. Zou, S. et al. Ag Nanorods-Based Surface-Enhanced Raman Scattering: Synthesis, Quantitative Analysis Strategies, and Applications. *Frontiers in Chemistry* 7 (2019).
 25. Barbillon, G., Graniel, O. & Bechelany, M. Assembled Au/ZnO Nano-Urchins for SERS Sensing of the Pesticide Thiram. *Nanomaterials* 11, 2174 (2021).

26. Seo, M. & Ha, J.W. Effective surface-enhanced Raman scattering of randomly branched gold nano-urchins with Rhodamine 6G as Raman reporters. *Microchemical Journal* 140, 47 (2018).
27. Guo, P. et al. Plasmonic core-shell nanoparticles for SERS detection of the pesticide thiram: size- and shape-dependent Raman enhancement. *Nanoscale* 7, 2862 (2015).
28. Kumari, G. & Narayana, C. New Nano Architecture for SERS Applications. *The Journal of Physical Chemistry Letters* 3, 1130 (2012).
29. Lee, P.C. & Meisel, D. Adsorption and surface-enhanced Raman of dyes on silver and gold sols. *The Journal of Physical Chemistry* 86, 3391 (1982).
30. Creighton, J.A., Blatchford, C.G. & Albrecht, M.G. Plasma resonance enhancement of Raman scattering by pyridine adsorbed on silver or gold sol particles of size comparable to the excitation wavelength. *Journal of the Chemical Society, Faraday Transactions 2: Molecular and Chemical Physics* 75, 790 (1979).
31. Gibson, K.F., Correia-Ledo, D., Couture, M., Graham, D. & Masson, J.-F. Correlated AFM and SERS imaging of the transition from nanotriangle to nanohole arrays. *Chemical Communications* 47, 3404 (2011).
32. Zhao, X. et al. Design of Hybrid Nanostructural Arrays to Manipulate SERS-Active Substrates by Nanosphere Lithography. *ACS Applied Materials & Interfaces* 9, 7710 (2017).
33. Wang, J.F., Wu, X.Z., Xiao, R., Dong, P.T. & Wang, C.G. Performance-Enhancing Methods for Au Film over Nanosphere Surface-Enhanced Raman Scattering Substrate and Melamine Detection Application. *PLOS ONE* 9, e97976 (2014).
34. Karadan, P., Aggarwal, S., Anappara, A.A., Narayana, C. & Barshilia, H.C. Tailored periodic Si nanopillar based architectures as highly sensitive universal SERS biosensing platform. *Sensors and Actuators B: Chemical* 254, 264 (2018).
35. Uetsuki, K., Verma, P., Nordlander, P., Kawata, S. Tunable plasmon resonances in a metallic nanotip-film system. *Nanoscale* 4, 5931 (2012).
36. Long, J., Yi, H., Li, H., Lei, Z. & Yang, T. Reproducible Ultrahigh SERS Enhancement in Single Deterministic Hotspots Using Nanosphere-Plane Antennas Under Radially Polarized Excitation. *Scientific Reports* 6, 33218 (2016).

37. Chih-Feng Wang, Z.C., Brian T. O’Callahan, Kevin T. Crampton, Matthew R. Jones, and Patrick Z. El-Khoury. Tip-Enhanced Multipolar Raman Scattering. *Journal of Physical Chemistry Letters* 11, 2464 (2020).
38. Israelsen, N.D., Hanson, C. & Vargis, E. Nanoparticle Properties and Synthesis Effects on Surface-Enhanced Raman Scattering Enhancement Factor: An Introduction. *The Scientific World Journal* 2015, 124582 (2015).
39. Vlasov, A.V. et al. Raman Scattering: From Structural Biology to Medical Applications. *Crystals* 10 (2020).
40. Ahlinder, L., Wiklund Lindström, S., Lejon, C., Geladi, P. & Österlund, L. Noise Removal with Maintained Spatial Resolution in Raman Images of Cells Exposed to Submicron Polystyrene Particles. *Nanomaterials* 6 (2016).
41. Niaura, G. Raman Spectroscopy in Analysis of Biomolecules. *Encyclopedia of Analytical Chemistry*, 1 (2014).
42. Ronningen, T.J., Schuetter, J. M., Wightman, J. L., Murdock, A., Bartko, A. P. in *Biological Identification* (ed. Schaudies, R.P.) 313 (Woodhead Publishing, 2014).
43. Shipp, D.W., Sinjab, F. & Notingher, I. Raman spectroscopy: techniques and applications in the life sciences. *Advances in Optics and Photonics* 9, 315 (2017).
44. Almealmadi, L.M., Curley, S.M., Tokranova, N.A., Tenenbaum, S.A. & Lednev, I.K. Surface Enhanced Raman Spectroscopy for Single Molecule Protein Detection. *Scientific Reports* 9, 12356 (2019).
45. Yang, Y. et al. Roadmap for single-molecule surface-enhanced Raman spectroscopy. *Advanced Photonics* 2, 1 (2020).
46. Xu, G., Song, P. & Xia, L. Examples in the detection of heavy metal ions based on surface-enhanced Raman scattering spectroscopy. *Nanophotonics* 10, 4419 (2021).
47. Yakubovskaya, E., Zaliznyak, T., Martínez Martínez, J. & Taylor, G.T. Tear Down the Fluorescent Curtain: A New Fluorescence Suppression Method for Raman Microspectroscopic Analyses. *Scientific Reports* 9, 15785 (2019).
48. Nunes, A., Magalhães, S. in *Raman Spectroscopy* (ed. Nascimento, G.M.d.) (IntechOpen, 2018).
49. Barton, S.J. & Hennelly, B.M. An Algorithm for the Removal of Cosmic Ray Artifacts in Spectral Data Sets. *Applied Spectroscopy* 73, 893 (2019).

50. Ramírez-Elías, M.G., & González, F. J. in Raman Spectroscopy (ed. Nascimento, G.M.d.) (IntechOpen, 2018).
51. Bumbrah, G.S. & Sharma, R.M. Raman spectroscopy – Basic principle, instrumentation and selected applications for the characterization of drugs of abuse. *Egyptian Journal of Forensic Sciences* 6, 209 (2016).
52. Jones, R.R., Hooper, D.C., Zhang, L., Wolverson, D. & Valev, V.K. Raman Techniques: Fundamentals and Frontiers. *Nanoscale Research Letters* 14, 231 (2019).
53. Siddhanta, S., Karthigeyan, D., Kundu, P.P., Kundu, T.K. & Narayana, C. Surface enhanced Raman spectroscopy of Aurora kinases: direct, ultrasensitive detection of autophosphorylation. *RSC Advances* 3, 4221 (2013).
54. Aggarwal, S. et al. Divalent Ion-Induced Switch in DNA Cleavage of KpnI Endonuclease Probed through Surface-Enhanced Raman Spectroscopy. *The Journal of Physical Chemistry B* 125, 2241 (2021).
55. Narayana, C., Aggarwal, S. Tailored plasmonic nano architectures for surface enhanced Raman spectroscopy and its applications in probing biomolecular interactions. *Ph.D thesis, Jawaharlal Nehru Centre for Advanced Scientific Research, Bengaluru*, 103 (2020).
56. Ghosh, S. et al. Structural heterogeneity in biliverdin modulates spectral properties of Sanderocyanin fluorescent protein. *bioRxiv*, 2021.04.02.438172 (2021).
57. Mantelingu, K. et al. Activation of p300 Histone Acetyltransferase by Small Molecules Altering Enzyme Structure: Probed by Surface-Enhanced Raman Spectroscopy. *The Journal of Physical Chemistry B* 111, 4527 (2007).
58. Pavan Kumar, G.V., Ashok Reddy, B.A., Arif, M., Kundu, T.K. & Narayana, C. Surface-Enhanced Raman Scattering Studies of Human Transcriptional Coactivator p300. *The Journal of Physical Chemistry B* 110, 16787 (2006).
59. Narayana, C., Pavan Kumar G.V. Utilization of surface enhanced Raman scattering in biomolecular detection and characterization. *Ph.D thesis, Jawaharlal Nehru Centre for Advanced Scientific Research, Bengaluru*, 113 (2008).

Chapter 2. Experimental Section

In this thesis, Raman spectroscopy and SERS have been used extensively to study lysozyme aggregation, detect extracellular vesicles, and design strategies for detecting mutations in DNA. This chapter discusses the various tools required for characterisations like UV-Visible spectroscopy, dynamic light scattering and zeta potential measurements, and transmission electron microscopy. Along with these, the methodology for experimental techniques like Raman spectroscopy and polymerase chain reaction are described. The details of the synthesis and characterisation of silver and gold nanoparticles required for SERS studies are provided.

2.1. UV-Visible Absorption

The UV-Visible absorption studies have been conducted on Lambda 750 UV/Vis/NIR spectrometer (Perkin Elmer). It works on the principle of attenuation of a beam of light when it passes through a sample as seen in Fig. 2.1.

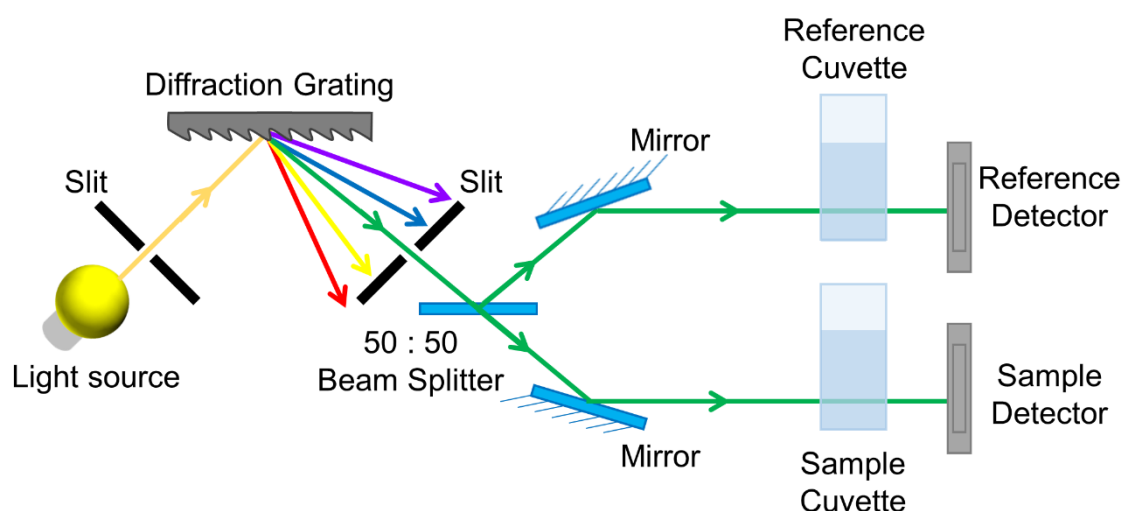


Fig. 2.1. Schematic of the UV-Visible absorption spectroscopy.

The absorption spectra were collected for all the samples in the 250–800 nm wavelength range. The peak position and FWHM for the spectra of the silver and gold nanoparticles were used to determine their approximate size distribution.

2.2. Dynamic Light Scattering and Zeta Potential

Dynamic light scattering works on the principle of light scattering from particles in a suspension undergoing the Brownian motion. Light is collected at an angle, away from incidence angle as seen in Fig. 2.2. The hydrodynamic radius and the zeta potential of the sample were measured in the Zetasizer Ultra (Malvern Instruments) using a disposable plastic cell (DTS0012) for the DLS measurement and a disposable folded capillary Zeta cell (DTS1080) for the Zeta potential measurements at temperature 25°C.

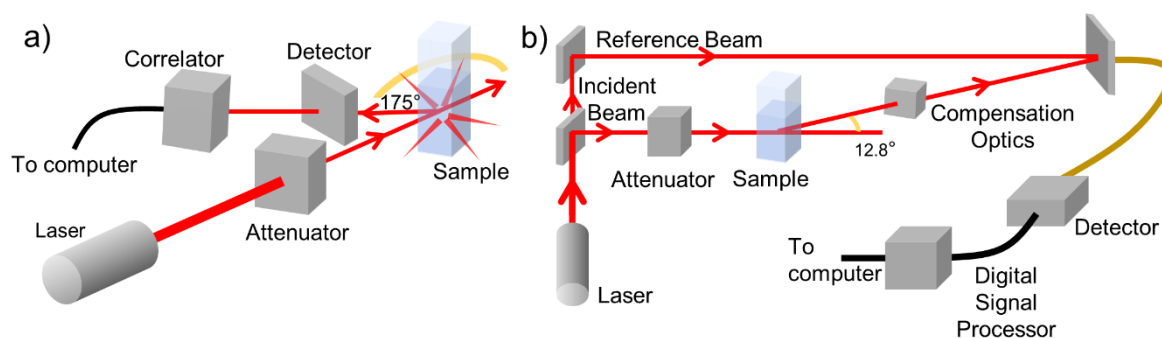


Fig. 2.2. Schematic of the DLS and the Zeta potential measurement modes in the Zetasizer.

2.3. Transmission Electron Microscopy (TEM)

The TEM images were obtained using a JEOL 300 TEM at an operating voltage of 200 kV for nanoparticles and 120 kV for EVs and proteins. Carbon-coated copper grids/ formvar (TEM-FCF200CU, Sigma Aldrich) were used for all imaging studies. For obtaining nanoparticle images, 3–5 μl of the dilute solution was dropped onto a grid and let dry. The TEM grid was prepared for the protein aggregation studies by dropping 3 μl of the protein solution and let dry for 5 min. The grid was washed thrice with water, and a filter paper was used each time to blot out the excess water. Then 2 μl of 2 % uranyl acetate was added, and the excess was removed using a filter paper after 30 s. For the images of the EVs, a small quantity of the diluted mixture was dropped onto a grid. When the sample was partially dry, a small amount of 1 % uranyl acetate stain was added, and after 4–5 min, the excess was blotted using a filter paper. For the EVs, cryo-TEM images were also obtained. The prepared grids were dried under an IR lamp and stored in a desiccator until the images were collected.

2.4. Raman Spectroscopy and SERS studies

The Raman spectroscopy studies were conducted on the LabRam HR Evolution (Horiba) spectrometer (Fig. 2.3). The system has an open-electrode CCD detector air-cooled to $-60\text{ }^{\circ}\text{C}$. Excitation lasers 532 nm and 633 nm with source powers of 100 mW and 17 mW, respectively, were used. Edge filters with steep cut-offs were used to obtain Raman spectra beginning from 50 cm^{-1} . The system was auto-calibrated at the beginning of the experiment using a reference silicon sample.

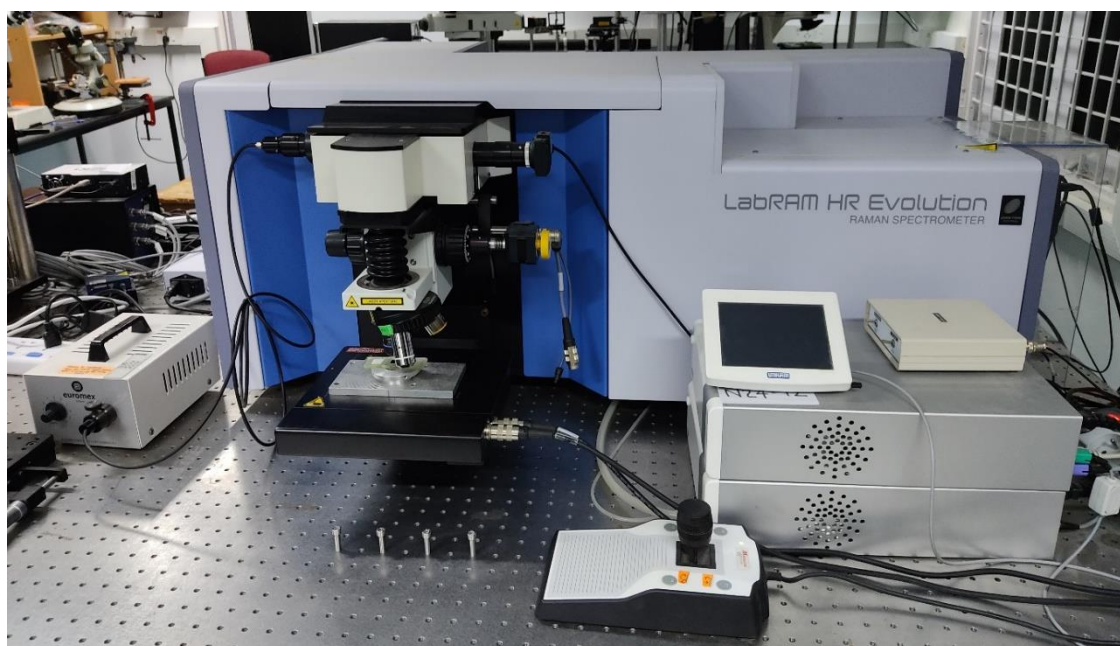


Fig. 2.3. LabRam HR Evolution Raman Spectrometer.

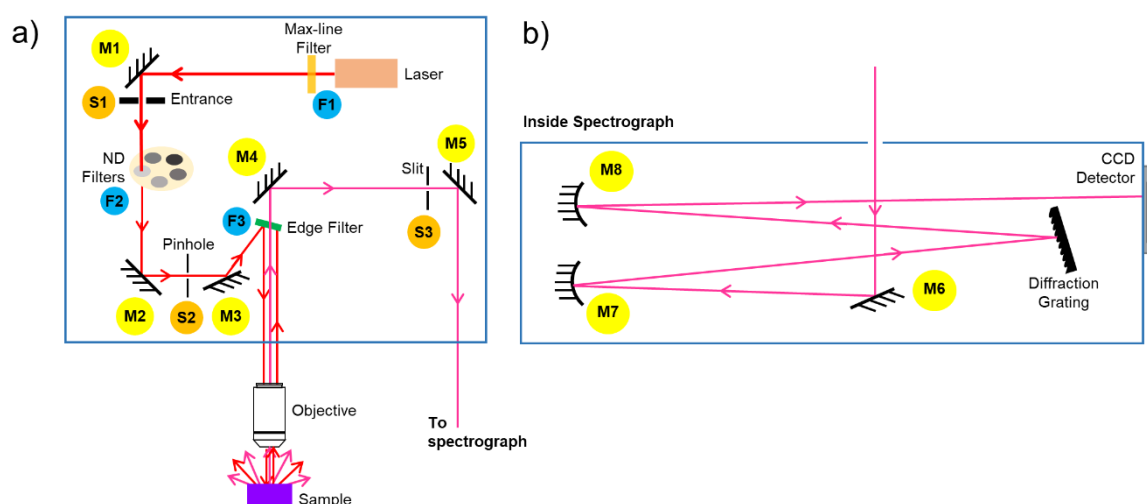


Fig. 2.4. a) Schematic of the optical path inside the LabRam HR Evolution, and b) Schematic of the optical path inside the spectrograph.

As seen in Fig. 2.4a, the light-emitting from a laser passes through a max-line filter (F1) to remove any laser glows or other unnecessary wavelengths of light. On reflection from a mirror (M1), it passes through an opening (S1) and a neutral density (ND) filter wheel (F2), where a specific ND filter can be chosen to help control the laser power that is being incident on the sample. On reflecting from a mirror M2 and passing through a pinhole (S2), the light is then incident on a mirror M3. The M3 is positioned so that the reflected beam falls on the edge filter (F3) at a specific angle. An edge filter is designed to reflect light below a cut-off wavelength and thus reflects the incident laser light. Through a series of mirrors, this light is focussed onto a sample through an objective. The position of the sample under the objective can be controlled using a joystick connected to the motorized XYZ stage. The optical image of the sample through the objective can be viewed on the computer with a camera attachment. This facilitates proper focussing and accurate choice of the spot on the sample for the laser to focus on. The scattered light from the sample is collected back by the objective, and the collimated beam of scattered light is incident on the edge filter. The edge filter reflects all the Rayleigh scattered light and transmits the Stokes Raman scattered rays (wavelength higher than the cut-off). These scattered rays are reflected from a mirror M4 and pass through the slit (S3).

This beam enters the spectrograph on reflection from the M5 mirror. Inside the spectrograph (Fig. 2.4b), the incident scattered light reflects from the mirror M6 and is incident onto a convex mirror M7. The reflected light is now incident on the diffraction grating. Either of the 600, 1200 or 1800 grooves/mm diffraction gratings can be chosen in this spectrometer. The scattered light gets separated into different rays at different angles depending on the wavelengths. The convex mirror M8 collimates the diffracted rays onto a CCD detector. The air-cooled CCD detector converts the incident photons into a readable Raman spectrum, readable on LabSpec 6 software. This software can be used for data processing and further analysis.

2.5. Polymerase Chain Reaction

Polymerase chain reaction (PCR) is a method used to amplify particular parts of a DNA using short DNA sequences called primers. The temperature of the sample is repeatedly raised and lowered to help a DNA replication enzyme copy the target DNA sequence. A custom-designed oligonucleotide sequence anneals to a complementary region in the template DNA. Then the DNA polymerase enzyme adds nucleotides to the 3' end of this oligonucleotide using it as a primer, thus generating an extended region of double-stranded DNA.

The PCR has three steps: denaturation, annealing and extension (Fig. 2.5).

- a) Denaturation: The DNA template is heated to 94 °C, breaking the weak hydrogen bonds that hold the DNA strands together in a helix. This allows for the DNA to open, creating separate single-stranded DNAs.
- b) Annealing: Depending on the primers and other oligonucleotides, the mixture is cooled anywhere from 50-70 °C. This allows the primers to bind (anneal) to their complementary strands in the template DNA.
- c) Extension: The reaction is then heated to 72 °C for the DNA polymerase to extend the primers, adding nucleotides onto the primer in sequential order, using the target DNA as the template.

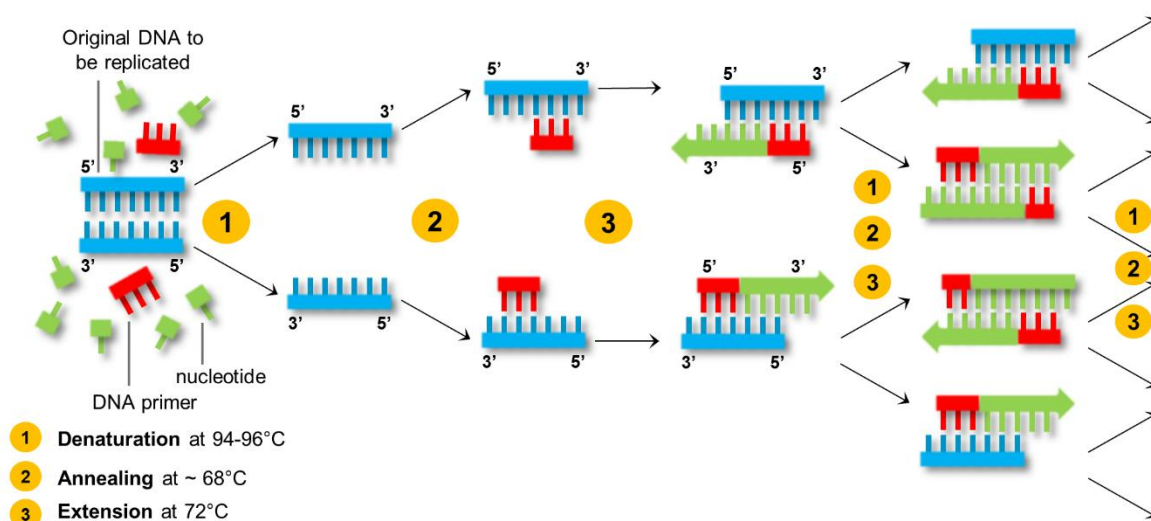


Fig. 2.5. Schematic representation of the PCR mechanism.

2.6. Synthesis of Silver Nanoparticles

The silver nanoparticles (AgNPs) required to conduct the SERS experiments were synthesized using the Lee-Meisel method (P.C Lee et al., *J. Phys. Chem*, 1982, 86, 3391-3395). In this method, sodium citrate reduces silver nitrate, resulting in citrate-capped silver nanoparticles (Fig. 2.6). Milli-Q water with a resistivity of 18.2 M Ω -cm at 25 °C was used. Silver nitrate and sodium citrate salts from Sigma-Aldrich were used for the synthesis.

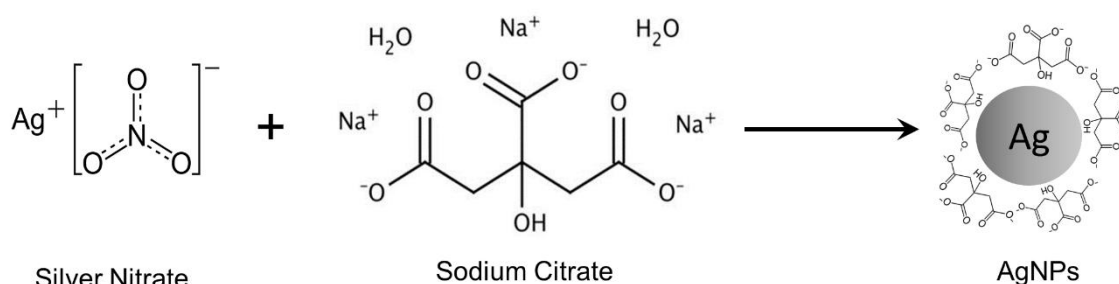


Fig. 2.6. Reaction mechanism for synthesis of silver nanoparticles by Lee-Meisel method.

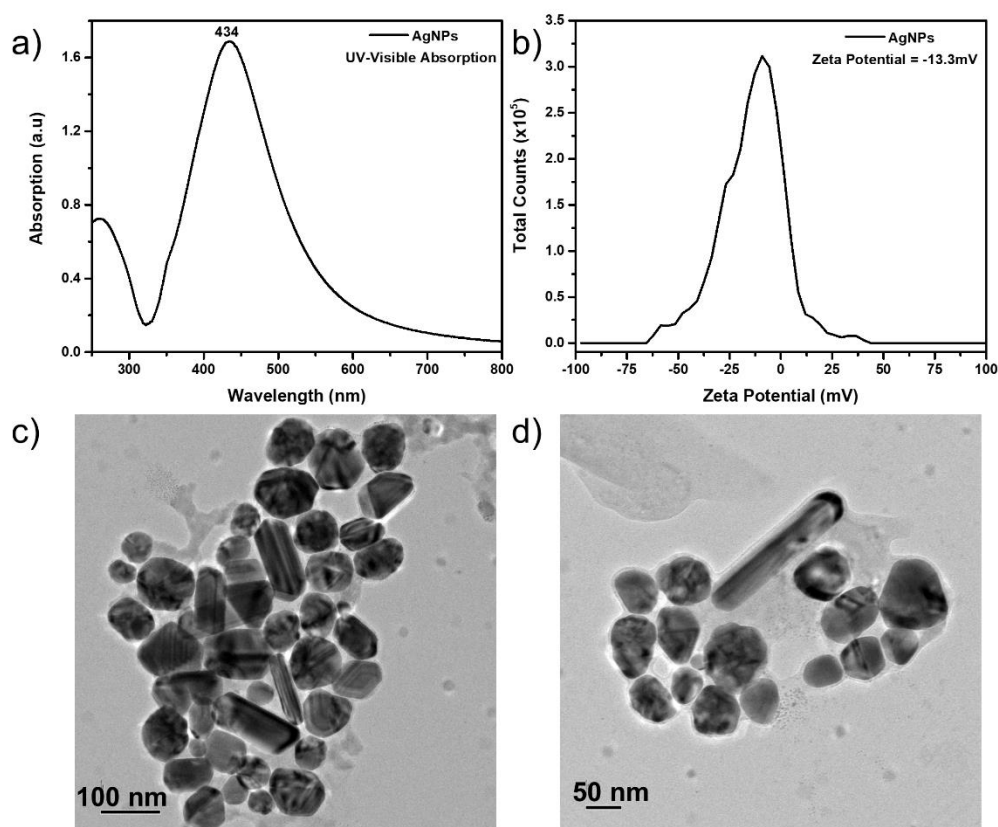


Fig. 2.7. a) UV-Vis absorption spectrum of the synthesized AgNPs, b) Zeta Potential, c) & d) TEM of the AgNPs.

The UV-Visible absorption spectra (Fig. 2.7a) of the synthesized AgNPs revealed a peak at 434 nm with an FWHM of ~ 120 nm, revealing a large size distribution. The zeta potential (Fig. 2.7b) of the AgNPs was -13.3 mV indicating negative citrate capping of the formed silver nanoparticles. TEM images (Fig. 2.7c&d) of the AgNPs, revealed a size distribution of ~ 20 – 120 nm.

2.7. Synthesis of Gold Nanoparticles

Various sizes of citrate-stabilized gold nanoparticles (AuNPs) were synthesized using the kinetically controlled seeded growth synthesis method proposed by Neus G. Bastús et al. (*Langmuir*, 2011, 27, 11098-11105). This method allowed for controlled growth of various sizes of AuNPs, with a higher concentration compared to standard procedures. Milli-Q water with a resistivity of 18.2 M Ω -cm at 25 °C was used for all the reactions. Sodium citrate and gold (III) chloride trihydrate salts from Sigma-Aldrich were used.

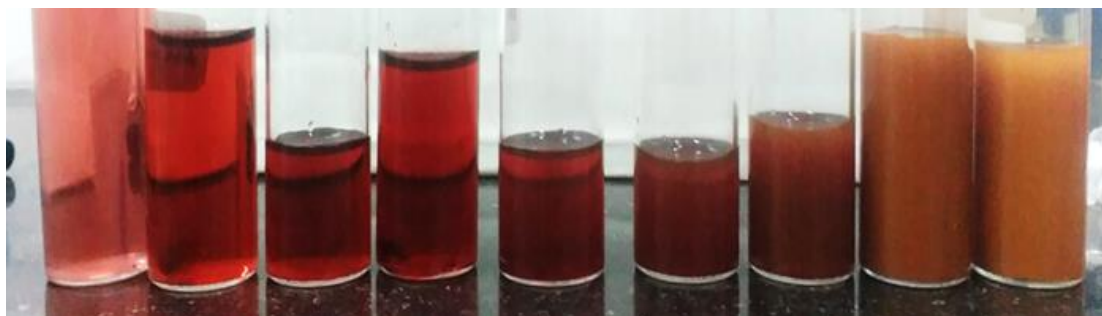


Fig. 2.8. Various stages of the synthesized AuNPs, show a variation in the colour and concentration. From left to right, the particle sizes and size distributions increase.

The procedure involved the synthesis of the Au seeds, followed by a controlled growth to obtain various sizes. A reflux mechanism was used to condense the water vapour and return it to the solution in the round bottom flask. To grow the seeds, 150 ml of 2.2 mM sodium citrate solution was brought to a vigorous boil with constant stirring. When 1 ml of 25 mM gold chloride solution was added, the solution turned light pink, indicating the formation of the Au seeds. The temperature of the solution was brought down to 90 °C after 10 min. 55 ml of the solution was removed, followed by the addition of 53 ml water and 2 ml of 60 mM sodium citrate solution. After 30 min, 1 ml of 25 mM gold chloride solution was added. After 30 min, 55 ml of the

solution was removed as stage 1 sample. These 2 steps of removal of the solution, followed by adding more precursors, were repeated multiple times to obtain multiple stages of the AuNPs (Fig. 2.8).

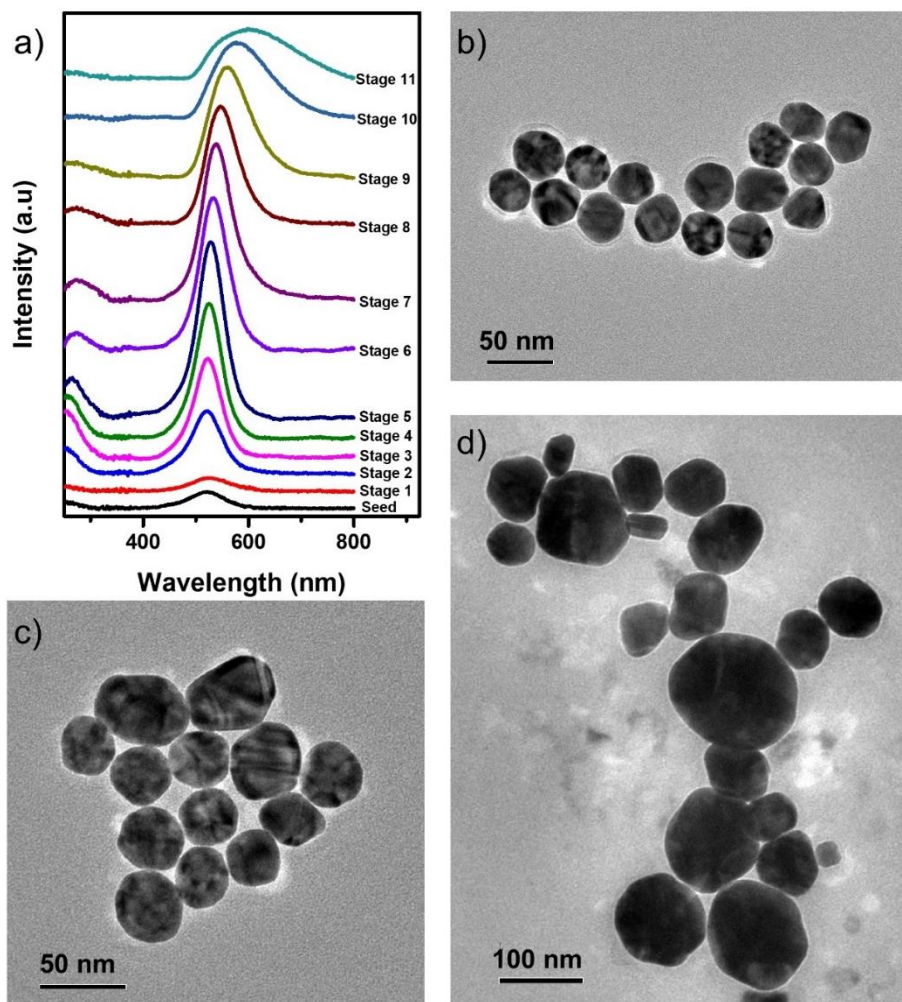
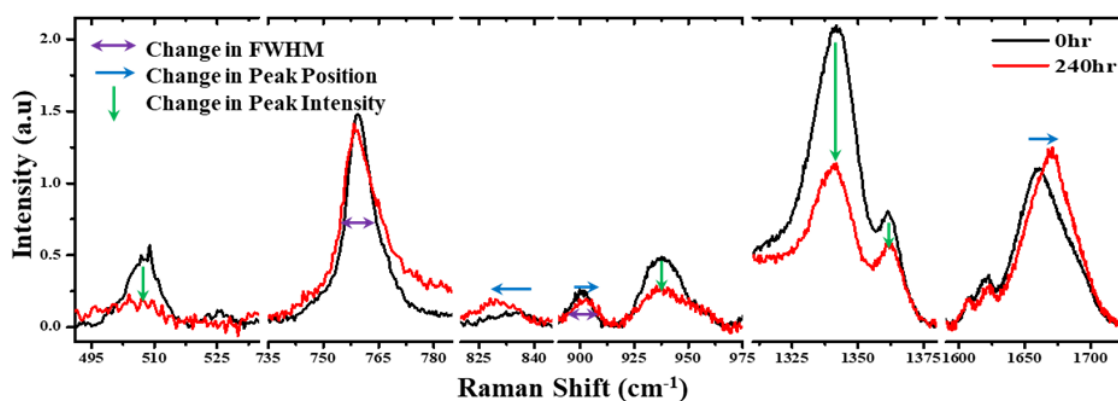


Fig. 2.9. a) UV-Visible absorption of the stages 1 to 11 of the synthesized AuNPs, TEM images of b) Stage 3, c) Stage 5, and d) Stage 8.

The UV-visible absorption spectra for the different AuNP stages were obtained, as seen in Fig. 2.9a. As the AuNP synthesis stages increase, the nanoparticles' size increases, as shown by the shift of the absorption peak to a higher wavelength. The increase in the FWHM over the stages indicates an increase in the size distribution, thus decreasing the mono-dispersion. TEM images of stage 3 and stage 5 AuNPs, as shown in Fig. 2.9b & c, show an almost monodisperse AuNPs of uniform size of ~ 30 nm and ~ 40 nm, respectively. The TEM image of the stage 8 AuNPs (Fig. 2.9d) shows an average ~ 70 nm sized particles, but polydisperse with a 30–130 nm size range.

Chapter 3. Understanding aggregation of Hen Egg White Lysozyme and Human Lysozyme in alkaline pH using Raman Spectroscopy and MD simulations*

Protein misfolding and aggregation play a significant role in the cause of neurodegenerative diseases like Alzheimer's, Parkinson's and so on. Though the aggregation of Hen Egg White Lysozyme (HEWL) has been studied from many different aspects using various experimental and computational techniques, each tends to reveal only one or a few properties from the study. This paper uses Raman spectroscopy, a versatile tool, to study multiple aspects of the protein aggregation process. The shift in the Amide I band from 1657 cm^{-1} to 1672 cm^{-1} reveals that, upon incubation in a pH 12.2 buffer at $25\text{ }^{\circ}\text{C}$, HEWL goes from an alpha-helical dominated system to a beta-sheet dominated structure, in ten days. Along with the inhibitor molecule studies, it can be interpreted that the protein aggregates by forming beta-sheets amidst the monomers. Molecular Dynamic simulations helped establish and visualize the unfolding of the proteins when exposed to an alkaline solution when the disulphide bonds are broken. With the help of protein docking, a preferential dimer formation was observed for the protein in pH 12.2 compared to the neutral pH system.



*This work has been carried out in collaboration with Dr. Amrendra Kumar and Prof. Swaminathan from IIT Guwahati & the theoretical calculations were performed by Pratik Behera and Dr. Shijulal Nelson from Rajiv Gandhi Centre for Biotechnology, Thiruvananthapuram.

3.1. Introduction

As we age, we tend to have higher chances of suffering from neurodegenerative diseases, most of which result from protein misfolding and aggregation. Sometimes genomic mutations, the functioning of a ribosome, or just a random event that occur can alter the folding pathway of the proteins upon their production.¹ These toxic configurations can interact with the existing copies and alter their state, hence are “infectious”.

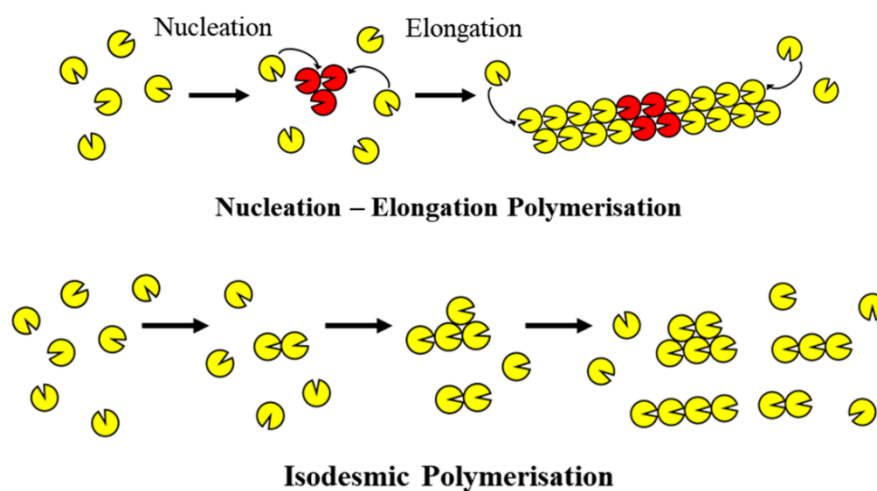


Fig. 3.1. Types of aggregation mechanisms, a) Nucleation-elongation polymerisation, and b) isodesmic polymerisation.

Protein aggregation can occur due to several unique mechanisms/pathways like the reversible association of the native monomer, aggregation of conformationally-altered monomer, aggregation of chemically-modified product, nucleation controlled aggregation and surface-induced aggregation.² Partially unfolded intermediates have dangling bonds, hence have exposed regions capable of intermolecular interactions.³ This complete or partial unfolding can be triggered by various physiochemical parameters, including temperature, pressure, pH, agitation, and chemical environment.⁴ The primary sequence of the protein along with the secondary and tertiary structures can play a role in the susceptibility of the protein to unfolding. Broadly, protein aggregation pathways can be classified into two standard models: nucleation-elongation polymerization and isodesmic polymerization. (Fig. 3.1) When aggregating by the isodesmic model, proteins become dangerously susceptible to aggregation as there is no critical concentration required for the process to begin.

3.1.1. Aggregation in HEWL

Hen egg-white lysozyme has been studied to aggregate under various external conditions like acidic or alkaline pH and the presence of chemicals like ethanol or guanidine hydrochloride. There are quite a few reports that explain the aggregation process of HEWL in acidic pH at higher temperatures than room temperature.⁴⁻⁶ The pH plays a vital role as it causes an effect on the protonation state of the charged amino acids and the C-terminal and N-terminal of the protein. The change in the charge distribution around the protein affects the salt bridges and hydrogen bonds, essential for a protein's strength and overall structure.⁷ In an acidic pH environment, the negative charge residues get neutralized, increasing repulsion amongst the monomers. Hence an external aid such as increased temperature or agitation is required to facilitate aggregation. However, the HEWL aggregation at alkaline pH can be achieved at ambient room temperature, avoiding those harsh conditions.⁸ Moreover, this aggregation process is directed by the isodesmic pathway. Hence, it can be studied at relatively lower concentrations.

Various biochemical analysis techniques, including Dynamic Light Scattering,⁹ Gel Electrophoresis,¹⁰ size-exclusion chromatography,¹¹⁻¹³ atomic force microscopy,^{5, 6, 8, 14} electron microscopy,¹⁵ have been used to indicate the formation and also estimate the size of protein aggregates. Other techniques like circular dichroism,¹⁶ fluorescence anisotropy,⁸ NMR,¹⁷ mass spectroscopy,^{18, 19} IR spectroscopy,²⁰ and Raman spectroscopy^{4-6, 21, 22} provide rather indirect information on molecular interactions, hence the formation of protein aggregates. Various forms of Raman spectroscopy, including Resonant Raman, Surface Enhanced Raman, FT-Raman, have been used to their advantage to be able to obtain specific information about biomolecular behaviours without getting lost amidst the vast information that exists in a Raman spectrum.²² Various components of a Raman spectrum like the peak position, peak shift, intensities, and FWHM can reveal information regarding various aspects of a protein and hence provide sensitive hints of unfolding and aggregation of the protein. Analysing an

obtained Raman spectrum can reveal many kinds of information about a system that would sum up the results of multiple different techniques making it a versatile tool.

3.1.2. Raman spectroscopy of proteins

Proteins are bulky molecules with hundreds of atoms, making the vibrational spectra complex with many normal modes. However, it is possible to interpret the different spectral regions separately. As the normal modes of all the repeating units occur at specific ranges, the Raman spectrum can be analysed. From interpreting these bands, information about the structure and environment of the amino acid side chains can be obtained. The band positions and line widths reveal information regarding the bound ligands and the overall protein backbone itself. Hence, Raman spectroscopy allows studying proteins of various sizes and in different environments.

Proteins constitute the amino acid chain backbone, the aromatic amino acids, disulphide bridges, and the total protein's net secondary structure. These various components have signature peaks in different parts of the fingerprint region of a Raman spectrum. Their analysis under different conditions can reveal changes happening to these components.

a) Amino acid chain backbone:

A linear sequence of the amino acids results in the protein's primary structure. The hydrogen bonding of the peptide backbone causes the amino acids to fold into a repeating pattern resulting in its secondary structure. Finally, the side-chain interactions lead to the three-dimensional folding of the protein, indicated by the tertiary structure. Certain proteins consist of more than one amino acid chain, causing a quaternary structure. The C and N atoms in the backbone are connected by single bonds that twist and rotate resulting in different vibrational modes (Fig. 3.2).

The CH₂ symmetric rocking, C_α-C stretching, CH₂ twisting and wagging, N-C_α-C stretch in the skeleton and CH₃, CH₂ and CH deformation and scissoring (1400-1500 cm⁻¹) are some of the unique vibrational modes corresponding to the backbone.

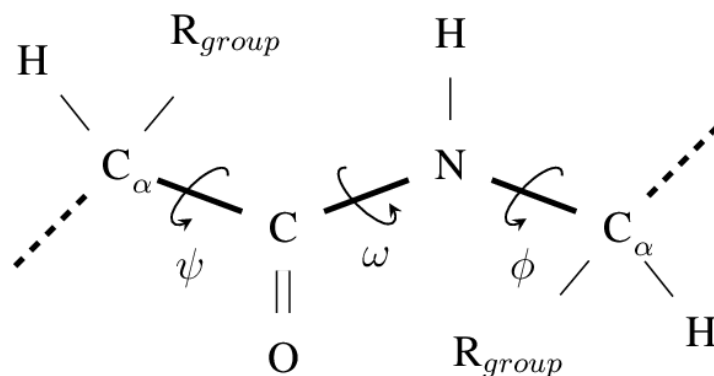


Fig. 3.2. Schematic representation of a polypeptide chain with two amino acids, showing average distances of bonds. The main chain atoms between two C_α atoms are generally fixed in a plane, with the C-N peptide bond angle (ω) being 0° or 180° (trans, most common, or cis forms), thus only the C_α -C and N- C_α bonds exhibit rotational mobility.

b) Aromatic amino acids:

The aromatic rings of the aromatic amino acids, phenylalanine, tyrosine, and tryptophan (Fig. 3.3) are highly polarisable and produce characteristic modes that can be used as Raman markers.

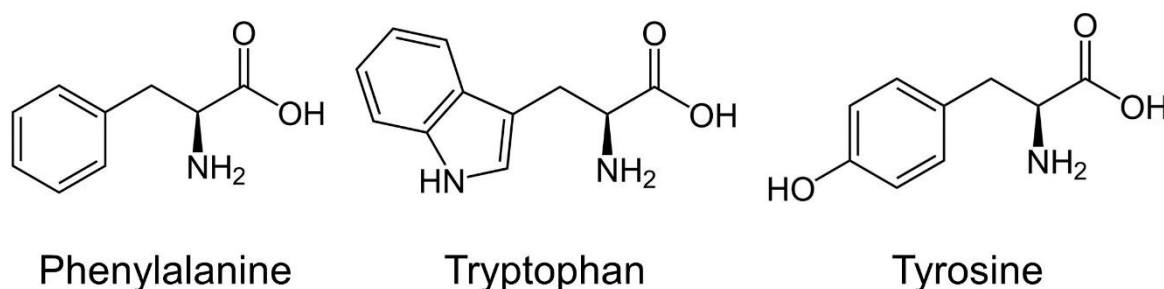


Fig. 3.3. The aromatic amino acids – phenylalanine, tryptophan and tyrosine.

The ring vibration of the phenylalanine produces a strong Raman peak around 1004 cm^{-1} . This vibration is generally not sensitive, hence unaffected by any conformational changes in the protein. So, this peak can be used as a reference peak to normalize the protein's Raman spectrum. Peak normalized spectra can be compared to comment on the intensity changes of the other peaks.

The indole ring vibrations contribute to the highest number of peaks from tryptophan. The relative intensities of the peaks corresponding to a fermi doublet at 1340 and 1360 cm^{-1} act as a hydrophobicity marker. If the intensity ratio of this doublet is high, then the tryptophan is exposed to the solvent and hence is in a hydrophilic

environment. Similarly, if the intensity ratio is low, it is buried inside the protein's hydrophobic environment. The band around 1010 cm^{-1} is sensitive to the van der Waals interaction of the phenyl ring. The higher its peak position, the greater is the interaction force.

There are two peaks at 830 and 850 cm^{-1} that constitute tyrosine's fermi doublet. They arise from the ring breathing fundamental and the overtone of the C–C–O deformation in the para-substituted benzene rings. The relative intensities of these peaks tell us the hydrogen bonding condition of the phenol group of the tyrosine residue.

c) Disulphide bridges:

The amino acid cysteine produces significant side-chain vibrations. Intra- or intermolecular disulphide bonds link cysteines in a protein as seen in Fig. 3.4. The S–S stretching vibrations produce Raman bands in the range of 500 to 540 cm^{-1} . Depending on the conformation of –C–S–S–C– to be gauche-gauche-gauche, gauche-gauche-trans, etc, their band positions vary. The C–S stretching frequency lies around 700 or 740 cm^{-1} , depending on whether the conformation is gauche or trans. The intensities of these peaks indicate the formation or breaking of disulphide bonds.

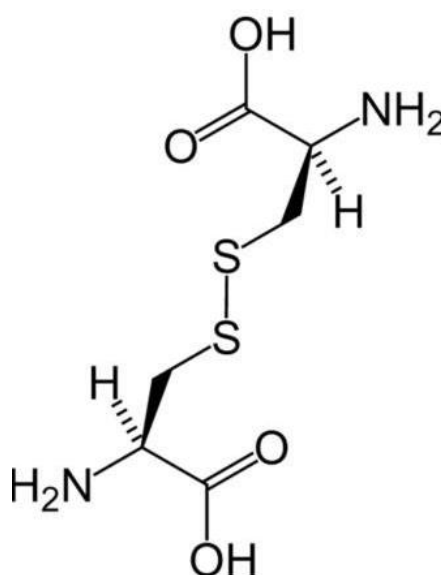


Fig. 3.4. The disulphide bond between two cysteines in a protein.

d) Secondary structure of the protein:

The secondary structure includes the alpha helices, parallel and anti-parallel beta-sheets, beta turns and the random or unordered coils (Fig. 3.5). These result from the hydrogen bonding between the various amino acids in the protein's primary sequence.

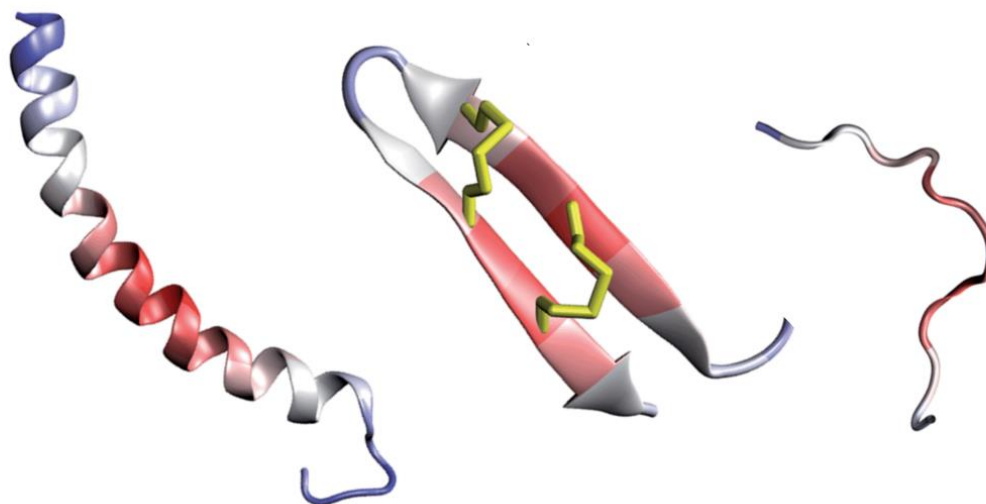


Fig. 3.5. Secondary structures in a protein – alpha helix, antiparallel beta-sheet with a beta-turn and a random coil (from left to right). (adapted from ²³)

The spatial arrangement of the C=O bonds and the coupling between the individual vibrations in a protein's secondary structure results in the amide I Raman band. The amide I, II and III are combination modes of N-H and C=O vibrations in a protein's secondary structure (Fig. 3.6). The amide I band's contribution is mainly from the C=O stretching (approx. 80%) and appears around 1650 cm^{-1} . The amide III band appearing around 1300 cm^{-1} has contributions from about 40% from the C-N stretch and 30% from the N-H bend. The amide II band originating from about 60% N-H bend and 40% C-N stretch, around 1550 cm^{-1} , is a weak band and doesn't appear in the absence of resonance. As the amide III band has contributions from both secondary structure and side chains, and amide I band only from the secondary structure, the amide I band is more suitable for secondary structure analysis.

The secondary structures comprising alpha-helices, parallel and anti-parallel beta-sheets, beta turns, and the random or unordered coils have unique amide I bands.

Hence the deconvolution of the amide I band of a protein predicts the fraction of each secondary structure from their relative intensities. The shape and position of this band are affected by various interactions.

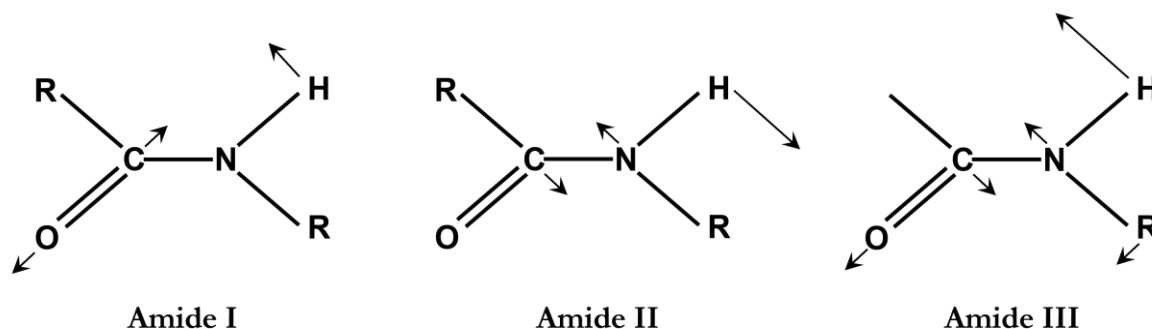


Fig. 3.6. Vibrations corresponding to the amide bands in a protein. (adapted from Foggia et al²⁴)

3.1.3. MD simulations for protein aggregation studies[†]

Over the past many years, with the advancement of tools and computer hardware, molecular dynamic simulations have achieved a leading part in understanding protein dynamics.²⁵⁻²⁸ Supercomputers like Anton are specifically designed to perform MD simulations. They can simulate from nanoseconds to seconds the time-scale most molecular activities take.²⁹ Despite all these advancements in the field, this technique is still underutilized for protein aggregation, with very few MD investigations so far. The studies that have been performed mainly concentrate on amyloid formation using coarse-grained protein models simulated with or without explicit solvents.³⁰ These methods provide an appropriate picture of how proteins aggregate, but the molecular level details are sacrificed. Simulations can provide an extraordinary amount of detail concerning individual particle motions as a function of time and answer specific questions about the properties of a model system more often than experiments on an actual system.³¹

In this chapter, Raman spectroscopy combined with Molecular Dynamics (MD) simulation is utilized to gain molecular insights into the aggregation process of HEWL in pH 12.2 at ambient room temperature. The Raman spectra of the HEWL protein

[†] Simulations provided by Dr. Shijulal's group, RGCB, Thiruvananthapuram

incubated in pH 7.0 buffer as a control and in pH 12.2 buffer as the experimental study are studied at various time points to understand their response to an environment over time. Simultaneously, MD simulations are used to visualize and understand protein unfolding and look at the dimerization of the protein through docking studies.

3.2. Materials and Methods

3.2.1. Materials

The hen egg-white lysozyme (HEWL) and recombinant human lysozyme (HL) expressed in rice were purchased from Sigma Aldrich. Milli-Q water (Millipore Ltd) with resistivity at 18.2 M Ω .cm at 25°C was used for all the experiments. Sodium dihydrogen phosphate (NaH₂PO₄), sodium hydrogen phosphate (Na₂HPO₄), sodium hydroxide (NaOH), sodium azide, dimethylformamide and iodoacetamide from Sigma Aldrich were used for preparing buffers and inhibition studies.

3.2.2. Experimental Methods

a) Lysozyme Sample Preparation and Incubation:

A 50 mM stock of sodium dihydrogen phosphate was prepared, and the pH was adjusted to 7.0 using 0.1 N NaOH. Similarly, a 50 mM stock of sodium hydrogen phosphate was prepared and adjusted to pH 12.2 using 0.1 N NaOH. 0.1% w/v Sodium azide was added to these buffer solutions to prevent any microbial growth. HEWL and HL stock (5 mg/ml) were freshly prepared using MilliQ water. The protein concentration was determined using UV-Vis absorption spectroscopy. 1 ml of the samples with the required concentration (120 μ M and 30 μ M) were prepared by diluting the stock using 50 mM pH 7.0 buffer for control experiments and 50 mM pH 12.2 buffer for aggregation experiments. The prepared samples were incubated in a 25 °C stackable incubator without agitation for ten days.

For the aggregation inhibition studies, iodoacetamide is dissolved in Dimethylformamide to obtain a 1.5 M stock. 2 μ l of this stock is added to the HEWL in pH 12.2 buffer, at the 2 hr time interval. Later 2 μ l is added similarly at 6 hr, 12 hr

and 24 hr time intervals, while the sample continued to incubate at 25 °C in an opaque centrifuge tube sealed with parafilm to prevent any escape or reaction of the iodoacetamide. At the end of 24hrs, 8 μl of the iodoacetamide stock solution is present in the aggregating HEWL sample (1 ml, 120 μM in pH 12.2 buffer), resulting in its final concentration of 12 mM. Similarly, a control inhibition experiment is designed by preparing the sample in a pH 7.0 buffer.

b) Sample preparation and Raman spectroscopy studies:

At every time point, small aliquots of 20 μl were removed from the primary sample vials to prepare the sample for Raman studies. 10 μl of the aliquoted sample was put onto a siliconized glass substrate, then placed inside a desiccator for 10 min to completely dry out the water from the sample (Fig. 3.7a). The drop-coated deposition Raman (DCDR) spectra³² of the samples were collected at room temperature by randomly scanning using an Olympus LMPlanFL 50X objective to focus the laser beam to an approximate diameter of 1 μm on the raised edge of the dried ring (Fig. 3.7b). Raman scattering, generated by ~ 6 mW power of the 532 nm laser, was collected for 120 s per window using an 1800 grooves/mm diffraction grating. The Raman spectra were recorded in the spectral range of 300 cm^{-1} to 1800 cm^{-1} .

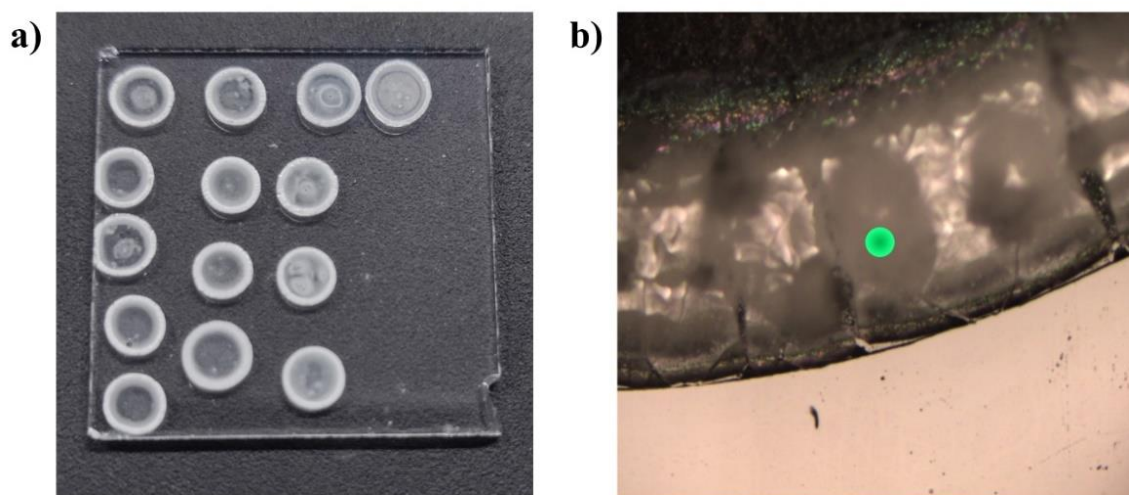


Fig. 3.7. a) Optical image of the rings formed when the sample is drop-casted onto a hydrophobic glass slide and dried in a desiccator, and b) 10X magnified image of the drop's edge from where the Raman spectra are collected

c) Data analysis (Baseline correction, normalisation and curve fitting):

Each Raman spectra was inspected on the LabSpec 6 software, and the cosmic radiation spikes were removed. A user-defined baseline was created and subtracted from the spectra on Origin Pro 2016 (OriginLab Ltd.). The spectra were then normalized with respect to the intensity of the phenylalanine band at 1004 cm⁻¹, as it is least affected by conformational changes.⁹ All the Raman peaks were fit using a Lorentzian curve. The amide I band around 1650 cm⁻¹ was deconvoluted into components fit by a Lorentz curve. The area under the deconvoluted peaks was calculated to form a direct reference to the percentage of a secondary structure component in the protein at a given time point. The time dependence of the various Raman peak parameters like peak position, FWHM or intensity was fit using a sigmoidal curve or exponential decay depending on growth or decay. A sigmoidal functional relationship indicates a lag phase, followed by a rapid growth phase and an equilibrium phase, and hence is ideal for this aggregation studies.

d) Dynamic light scattering (DLS) studies:

The hydrodynamic diameter of the aggregates was measured in the Zetasizer Ultra (Malvern Instruments) using a disposable plastic cell (DTS0012) at 25°C.

e) UV-visible absorption spectroscopy:

The prepared HEWL stock was diluted ten times using MilliQ water. The UV-visible absorption spectrum of this solution was obtained using Agilent 8453 UV-Visible spectrometer, and the absorption value at 280 nm was used to estimate the protein concentration of the stock with the extinction coefficient of 37970 M⁻¹cm⁻¹ using the following equation.³³

$$\text{Concentration of protein} = \frac{\text{absorption value at 280 nm} \times 10}{37970 \mu\text{M}} \quad (3.1)$$

3.3. Results and Discussion

The HEWL aggregation process was studied under alkaline conditions (pH 12.2), while the system under neutral conditions (pH 7.0) was used as the experimental control. The UV-Visible absorption spectrum of the HEWL stock was collected (Fig. 3.8), and the absorption value at 280 nm was used to calculate its concentration using an extinction coefficient of $37970 \text{ M}^{-1}\text{cm}^{-1}$. The absence of any peak around 320-350 nm confirms the absence of any aggregation states or UV-absorbing prosthetic groups in the system.³⁴ Raman spectrum was collected for the lyophilized monomer HEWL powder directly purchased from Sigma Aldrich as a reference. Raman spectra of $120 \mu\text{M}$ HEWL in pH 7.0 and pH 12.2 were collected at various time points, including 0, 0.5, 1, 3, 6, 12, 24, 48, 72, 96, 120 and 240 hr.

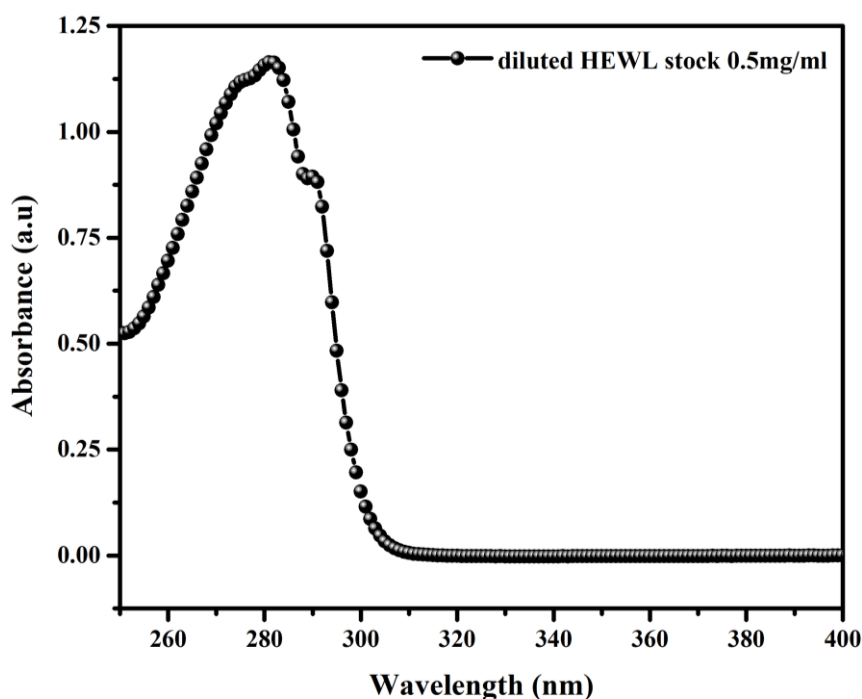


Fig. 3.8. UV-Visible absorption spectra of the diluted HEWL stock.

DCCR spectroscopy can be used to obtain Raman spectra by depositing the sample onto a substrate and letting it dry completely.³⁵ The obtained spectra had high reproducibility and were identical to those obtained from higher protein concentrations.^{32, 35} Using these results, we chose to conduct our experiments similarly and obtain DCCR spectra for all the time points without being concerned much about the loss of the native structure that would have been present in the solution form.

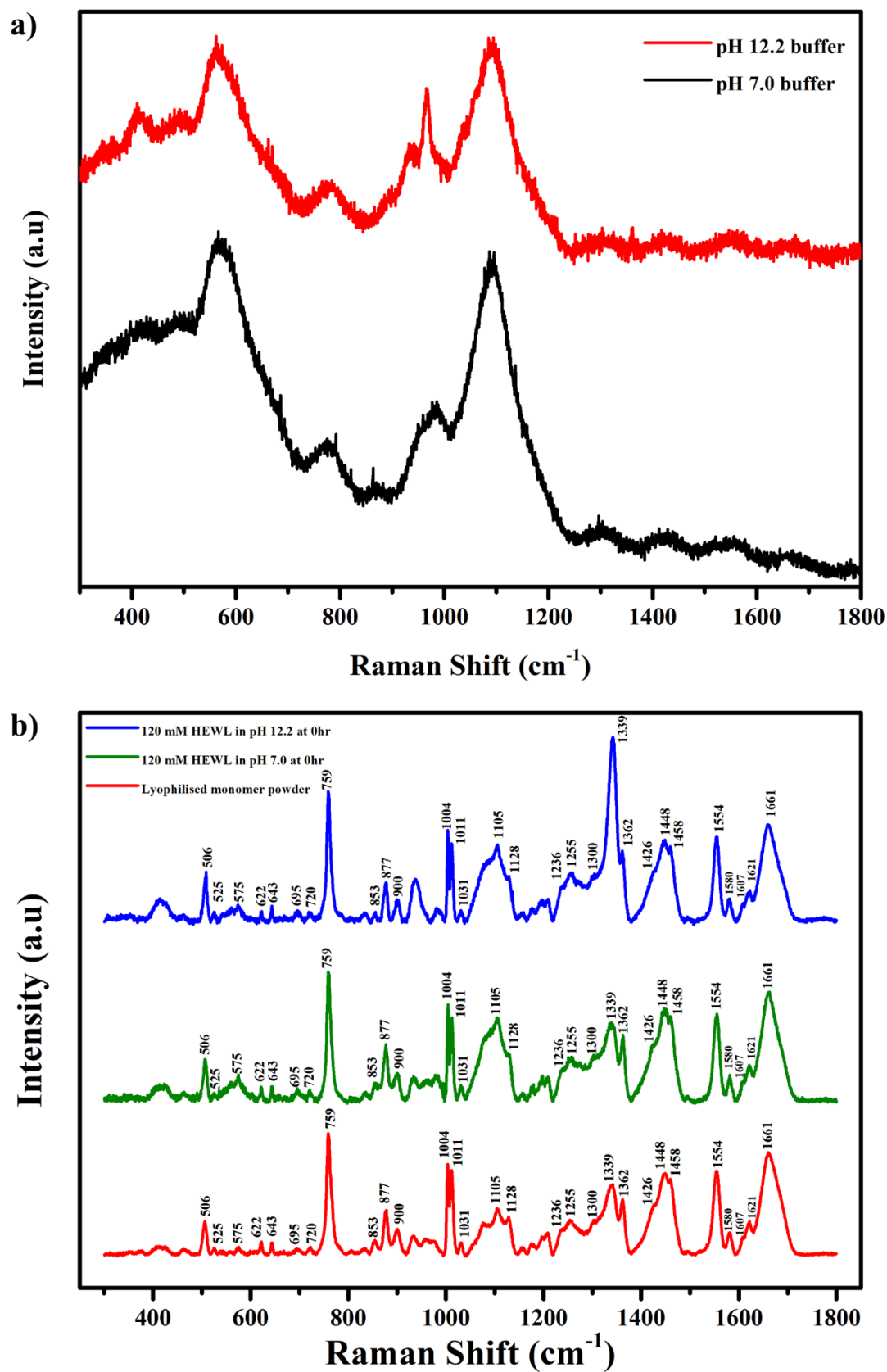


Fig. 3.9. a) Raman spectra of the buffers, and b) Raman spectra obtained from lyophilized HEWL monomer powder (red), ring edge of the 120 μM HEWL in pH 7.0 (green) and pH 12.0 (blue) at 0 hr of the experiment.

As a reference, the Raman spectra of the buffers were obtained to be sure that their contribution was subtracted before analysing the protein spectra (Fig. 3.9a). The Raman spectra in Fig. 3.9b shows all the peaks appearing in the dried control and sample compared to the monomer powder. The close similarity between the Raman profiles (Fig. 3.9b) obtained from lyophilized powder (red) and drop-casted sample (pH 7.0; green) indicates no/minimal artefacts in the acquired Raman signals from drop-casted samples. Thus, while using this methodology for sampling and signal acquisition, any change in the Raman profiles can directly be correlated to the changes in the chemical identity of the sample under investigation

A previous study by Ravi V K et al. gave insights into the isodesmic aggregation of HEWL caused by the formation of crosslinked disulphide bonds.⁸ Their steady-state fluorescence, DLS and AFM measurements proved that the HEWL aggregates formed irrespective of the initial concentration of the protein, thus making the process isodesmic. FRET was used to establish the formation of new intermolecular disulphide linkages, resulting in the formation of aggregate polymers. However, the precise change in the immediate molecular environment of interacting monomers leading to the observed aggregation is still unknown. Using MD simulations and Raman analysis, we tried to address this aspect.

First, the Raman spectrum of the protein was obtained, and the bands in the fingerprint region were assigned using previously studied literature. As this is a well-studied protein, past literature references were sufficient to assign the modes.

Table. 3.1. Peak assignments to the Raman modes obtained for HEWL.

Peak (cm ⁻¹)	Origin of Raman Mode	Ref.
505	S-S stretch (gauche-gauche-gauche) in cysteine	6, 36
525	S-S stretch (gauche-gauche-trans) in cysteine	6, 36
575	Tryptophan/cytosine	37
623	C-C twisting mode in Phenylalanine	37
643	C-C twisting mode of tyrosine	38
695	C-S bond	5
720	C-S stretch trans conformation	39
759	Tryptophan - Coupled vibrations of in-phase breathings of benzene and pyrrole (indole ring)	40
833 & 853	Fermi resonance of ring fundamental and overtone - Tyrosine	37, 39
877	Tryptophan - benzene ring + N ₁ H motion	22
900 & 930	N-C _α -C stretch of skeleton	5, 22
1004	Ring breathing of benzene in Phenylalanine	40
1011	Tryptophan ring breathing	37
1076, 1105, 1128	CH ₂ symmetric rock+ C _α -C stretching	5
1236	Type II β-turn strand, amide III	5
1255	Amide III	37
1300	CH ₂ - twisting, wagging	37
1339 & 1362	Tryptophan Fermi resonances between the fundamental in-plane N ₁ =C ₈ stretching and combination bands of ring out-of-plane deformations	40
1426	Pyrrole [$\nu(\text{N}_1\text{-C}_2\text{=C}_3) + \delta(\text{NH})$] + benzene $\delta(\text{CH})$	22
1448 & 1458	CH ₂ , CH ₃ and CH deformation and scissoring	5
1554	C=C Tryptophan	22, 37
1580	Tryptophan, amide II, Tyrosine	37
1607	Tyr, Phe ring vibration; C=C Phe, tyrosine	37
1621	$\nu_s(\text{ring}) + \delta(\text{OH})$ - Tryptophan	5, 22, 37
1661	Amide I band	5, 39

The plot represents these modes as seen in Fig. 3.10, and a detailed mode assignment is represented in Table. 3.1.

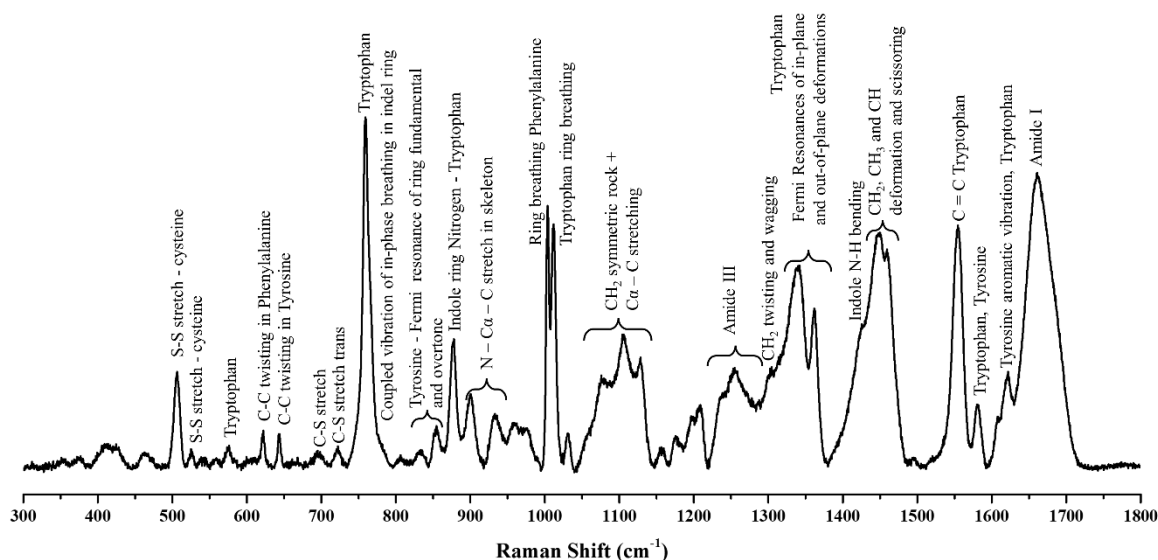


Fig. 3.10. Modes assigned to the various vibrations of HEWL in the Raman spectrum.

3.3.1. Unfolding of HEWL protein

The Raman peak around 505 cm^{-1} attributed to the S-S stretching is monitored over ten days at various time intervals. As seen in Fig. 3.11a, the peak remains constant for the protein in pH 7.0, while in pH 12.2, the peak starts to reduce at 12 hr time-point and almost completely disappears by the end of the 240 hr (Fig. 3.11b). The peak intensities of 505 cm^{-1} Raman peak were plotted with respect to time for the HEWL in both pH 7.0 and pH 12.2 buffers, which provided a clearer understanding of its behaviour (Fig. 3.11c). Over 240 hr, the peak intensity decreased, and FWHM increased. A similar observation was made in a study of HEWL aggregation in an acidic pH environment. This behaviour was assigned to the tertiary realignment of the core protein and/or hydrolysis of the disulphide bonds.⁵ This confirms that in an alkaline pH environment, the protein's disulphide bonds tend to break, thus resulting in the partial unfolding of the protein.

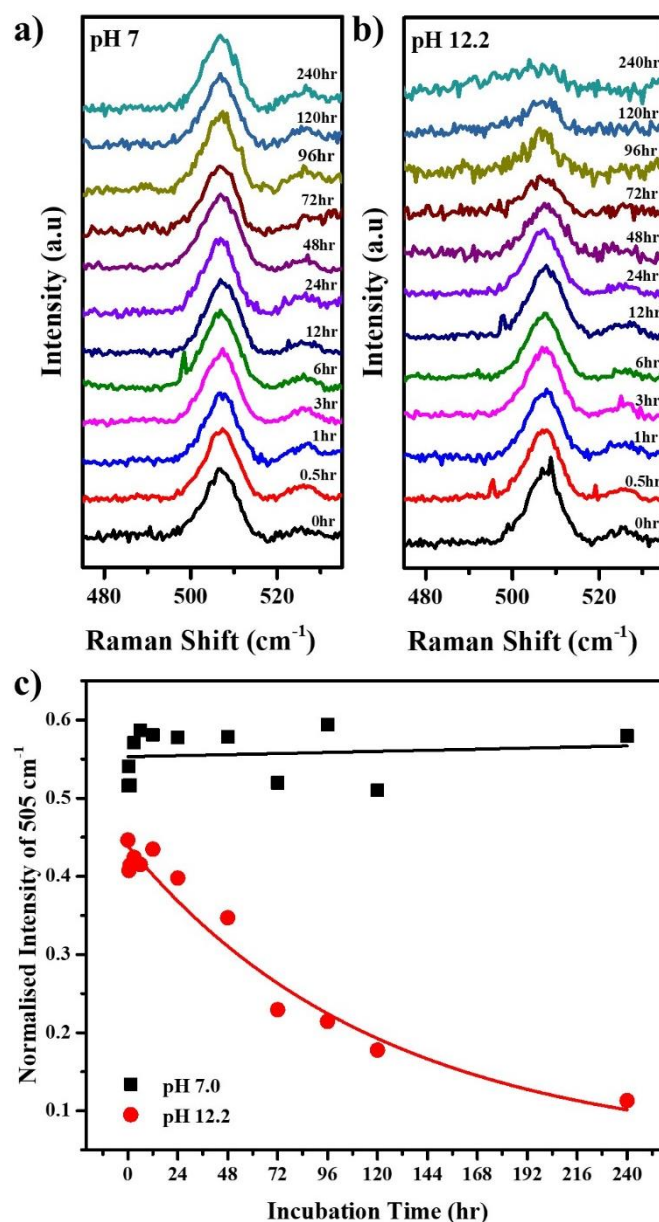


Fig. 3.11. Stack plots of 505 cm⁻¹ Raman peak in a) pH 7.0 and b) pH12.2 buffers at various time points of incubation; c) intensity vs time plot for 505 cm⁻¹ Raman peak.

As the correlation of real-time in MD simulations is quite challenging, studies have shown that higher environment temperatures hasten the unfolding process without affecting the unfolding pathway. The molecular dynamic simulations were conducted on the protein's 1HEW (PDB ID) crystal structure. For simulating the effect of pH, two systems were created, a) the protein monomer at neutral pH (pH 7.0) with all its disulphide bonds intact is the control system and is referred to as N, and b) the protein monomer at high pH (pH 12.2) with its intra-disulphide bonds broken is referred to as HBB. This was done to create an environment where the protein is

exposed to alkaline pH. When placed in pH 12.2, a protein undergoes an alkaline electrolytic shock, breaking many hydrogen bonds, salt bridges, and disulphide bonds,⁴¹ resulting in a secondary and tertiary structure loss.

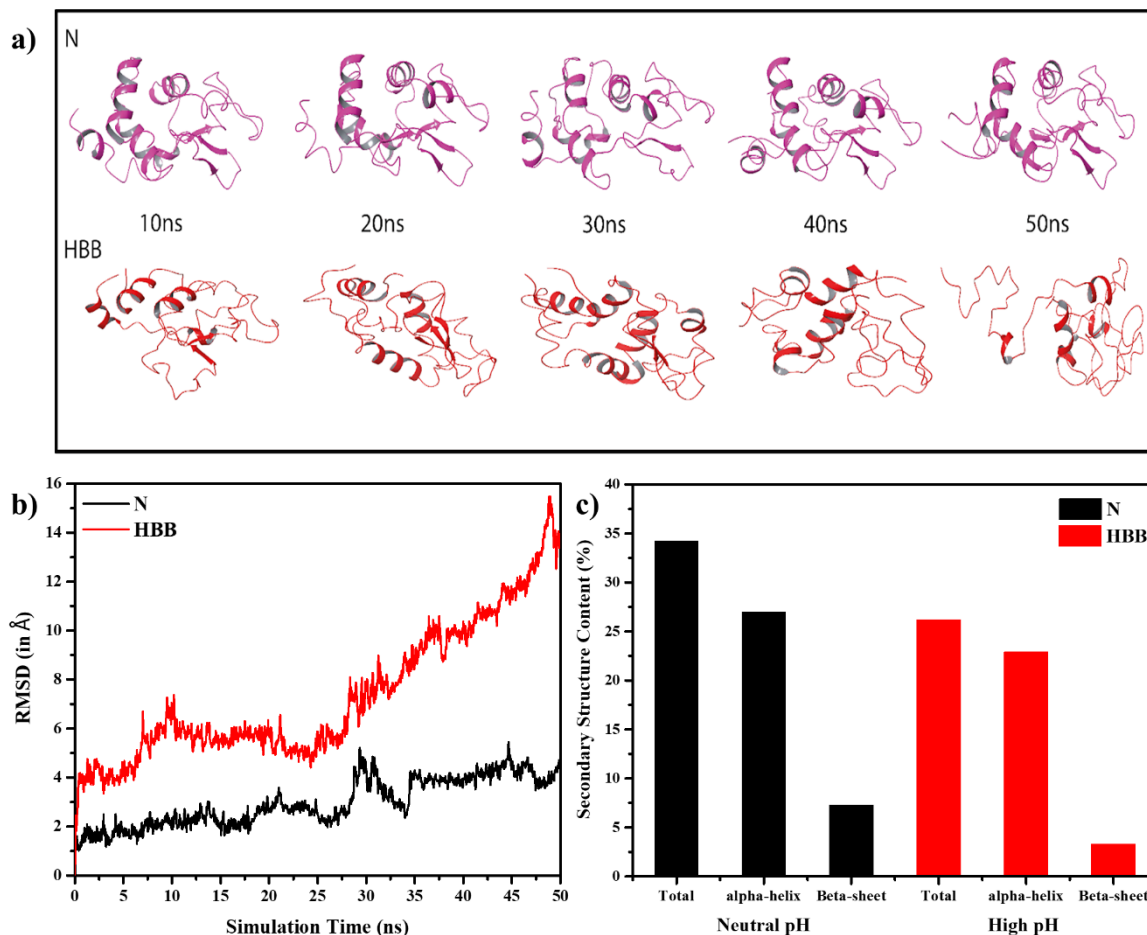


Fig. 3.12. a) Various frames of the HEWL monomer at different simulation time points, in N and HBB conditions; b) RMSD of the HEWL monomer in neutral pH and high pH with disulphide bonds broken in a 50 ns simulation; c) Secondary structure content in N and HBB after the simulation.[‡]

The monomers were subjected to 450 K for 50 ns each. This threshold of 450 K was defined to induce a restricted unfolding just enough to expose the protein surface for inter-monomeric interactions.⁴² The MD simulation trajectories comprising 1000 frames for 50 ns of independent simulations of monomers for structural and dynamic behaviour were examined, revealing further insights into the system. The backbone atoms computed through the 50 ns indicate that both the monomers retained a stable conformation until 25 ns. In N, there is a slight deviation in trajectory till 4 Å, whereas

[‡] Simulations performed by Dr. Shijulal's group from RGCB, Thiruvananthapuram

in HBB, the deviation in the trajectory reaches up to 16 Å, confirming that the system in alkaline pH becomes unstable when the disulphide bonds get broken (Fig. 3.12b). As seen in Fig. 3.12a, through the 50 ns of simulation, the secondary structure of N mostly remains intact, while HBB unfolds quite a bit. Through the simulation, the protein's secondary structure showed a decrease in alpha-helical content by about 5.97% and beta-sheet content by 4.4% in the HBB compared to the N is observed (Fig. 3.12c). Overall, the analysis of the 505 cm^{-1} peak in the Raman spectra and the MD simulations of the monomer in both conditions confirms that the disulphide bond breakage under alkaline conditions could be one of the major causes for the extensive unfolding of the protein leading to its higher propensity towards aggregation.

3.3.2. Protein aggregation and interactions

DLS measurements reveal the hydrodynamic radii of the samples, which gives us a preliminary idea of progress in the aggregation process of the system. On day 10 of the experiments, the hydrodynamic radius of the control in pH 7.0 buffer is around 2 nm (100%). The sample in pH 12.2 is distributed around 10.5 nm (30%) and 30.4 nm (70%) (Fig. 3.13). This increase in the hydrodynamic radius demonstrates that in alkaline pH, the protein aggregates, resulting in larger particles.

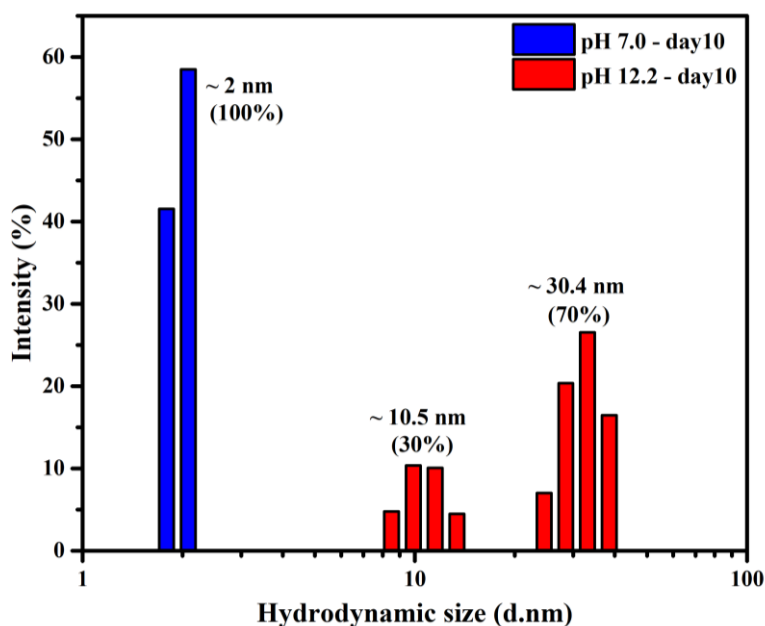


Fig. 3.13. DLS spectra of the protein samples on day 10 in pH 7.0 and pH 12.2 buffers.

The TEM images of the HEWL protein give a direct picture of the aggregation process, as seen in Fig. 3.14. It can be seen in the left image that the HEWL protein remained as scattered small particles when incubated in pH 7.0 for 10 days. As seen in the right image, longer polymeric chains of the protein have formed when incubated in the pH 12.2 buffer for 10 days, confirming aggregation.

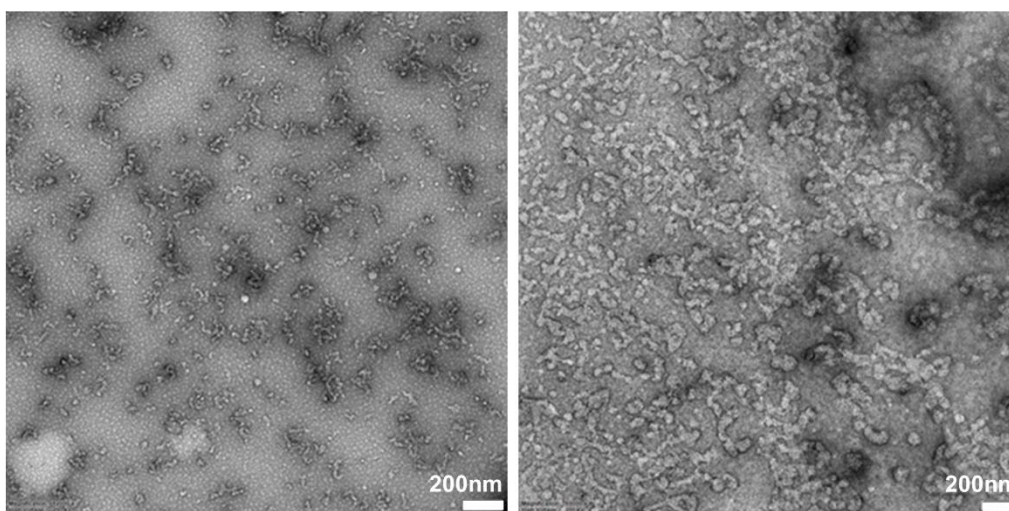


Fig. 3.14. TEM images of HEWL in pH 7.0 (left) and pH 12.2 (right) buffers on day 10.

3.3.2.1. Amide band analysis

The amide I band of HEWL in pH 12.2 showed a peak position shift over the incubation period (Fig. 3.15). The amide I band of the Raman spectrum of a protein provides information about the secondary structure of the protein.^{5, 9, 14, 40} The deconvolution of the amide I band (Fig. 3.16a) reveals its various constituents that include alpha helix (around 1657 cm^{-1}), beta-sheet (around 1671 cm^{-1}), beta turns (around 1688 cm^{-1}) and random coils/unordered protein (around 1647 cm^{-1}).

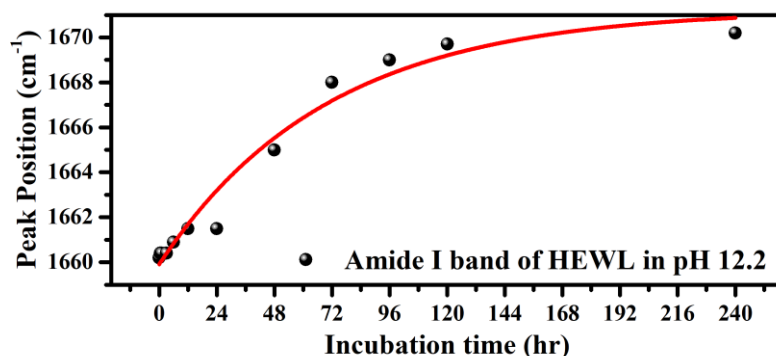


Fig. 3.15. Time evolution of the amide I band for HEWL incubated in pH 12.2 buffer.

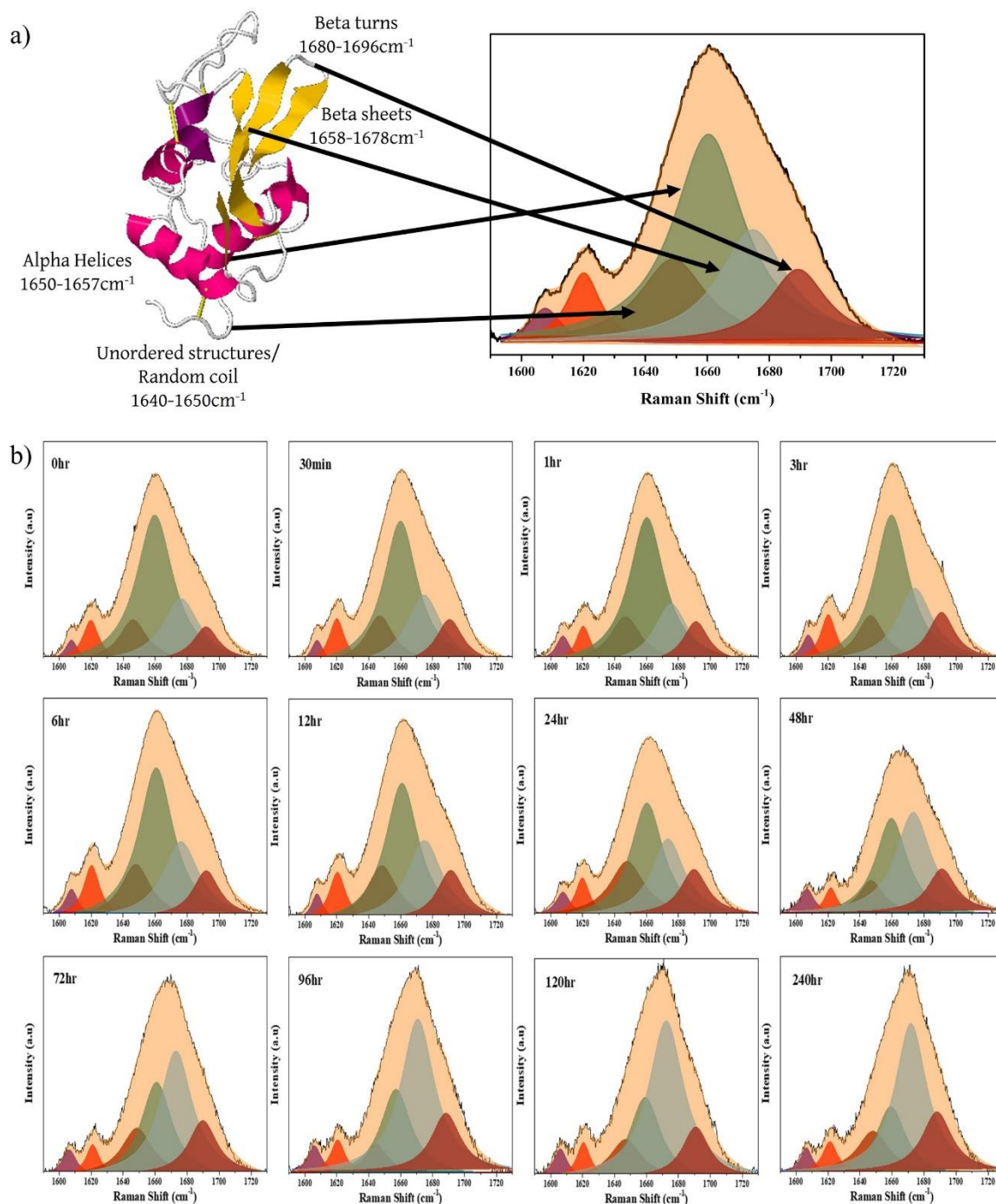


Fig. 3.16. a) Deconvolution of the amide I band into various secondary structure components, b) Deconvolution of the amide I band at various time points for the protein in pH 12.2. A gradual decrease in alpha-helical content and an increase in beta-sheet content after 12 hr of incubation of the protein in pH 12.2 is observed.

As shown in Fig. 3.17a, the position of the Amide I band of the protein in neutral pH remains constant. At the same time, a shift from an alpha-helix dominated system (peak around 1660 cm^{-1}) to a beta-sheet dominated system (peak around 1672 cm^{-1}) is observed in the alkaline pH buffer over time (Fig. 3.17b). Changes begin to

appear after a period of 24 hr. The amide I band was deconvoluted at every time point (Fig. 3.16b), and the percentage secondary structure of alpha-helix and beta-sheet was plotted as a function of time. It was observed that the alpha-helical and beta-sheet content remain almost similar, averaging around 50% and 23%, respectively, in the control system at pH 7.0 (Fig. 3.17c). But, in the system at pH 12.2 (Fig. 3.17d), the alpha-helical content decreased from about 50% to 20% over ten days. In comparison, the beta-sheet content increased from about 23% to about 45%, indicating the uncoiling of alpha helices and the formation of new beta-sheets structures.

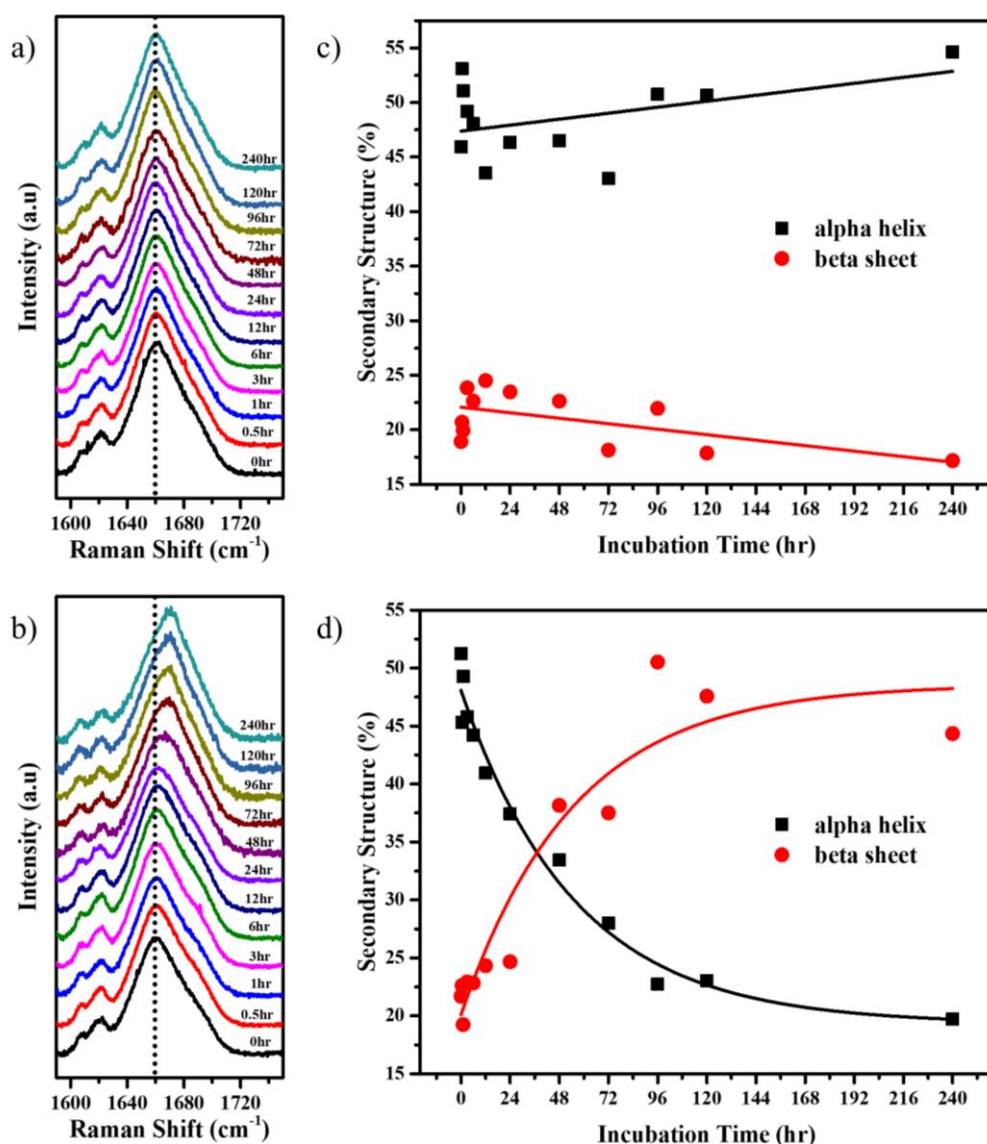


Fig. 3.17. Stack plot of the Amide I Raman band of HEWL in a) pH 7.0 and b) pH 12.2 buffers at various time points of incubation; Time evolution of the % alpha-helical and beta-sheet content of HEWL protein incubated in c) pH 7.0 and d) pH 12.2 buffers.

Isodesmic nature of the HEWL aggregation

Isodesmic polymerization's mechanism is independent of the initial concentration of the monomer. Unlike nuclear-elongation polymerization, it doesn't require a critical initial concentration for the process to begin. Depending on the initial concentration, here, different sizes of polymers are formed. The higher the concentration, the larger the size of the aggregate. To study the same, we used two initial concentrations of the HEWL protein, 30 μM and 120 μM .

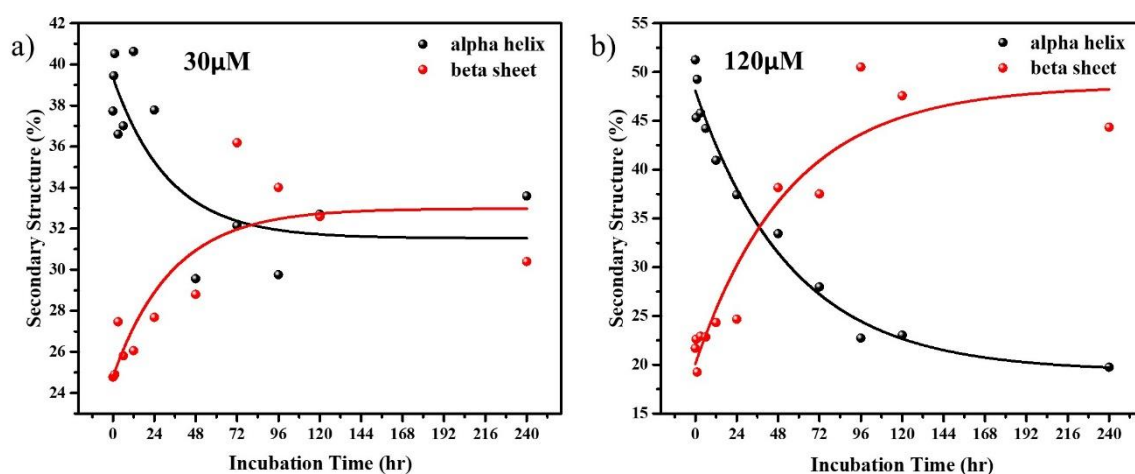


Fig. 3.18. Secondary structure analysis from amide I band for a) 30 μM and b) 120 μM HEWL.

Over the ten days of incubation in an alkaline pH 12.2 buffer, both the concentrations showed a shift in the amide I band. The amide I band was deconvoluted, and the total alpha-helical and beta-sheet content was extracted for the two concentrations of the HEWL protein. As seen in Fig. 3.18a and Fig. 3.18b, it is evident that in the lower 30 μM concentration, there is a lower percentage of beta-sheet content indicating smaller aggregates, while the higher percentage in the 120 μM HEWL protein indicates larger aggregates. The Raman studies performed on 3 μM of the HEWL protein was unsuccessful due to the poor Raman signal obtained at such low concentrations. As the amide band is one of the most intense bands, it was easy to obtain secondary structure analysis for both 30 and 120 μM of the HEWL protein. But most other bands for 30 μM were not so intense and hence challenging to analyse as a function of time. Therefore, the spectral analysis for the remaining peaks is demonstrated for the 120 μM HEWL protein.

The Amide III band around 1200-1300 cm^{-1} also gives insight into the secondary structure of the protein. The decrease in peak intensity around 1280 cm^{-1} and an increase in the peak around 1235 cm^{-1} also confirm the fact obtained from the amide I band analysis (Fig. 3.19) by indicating a decreased alpha-helical content and increased beta-sheet content, respectively.

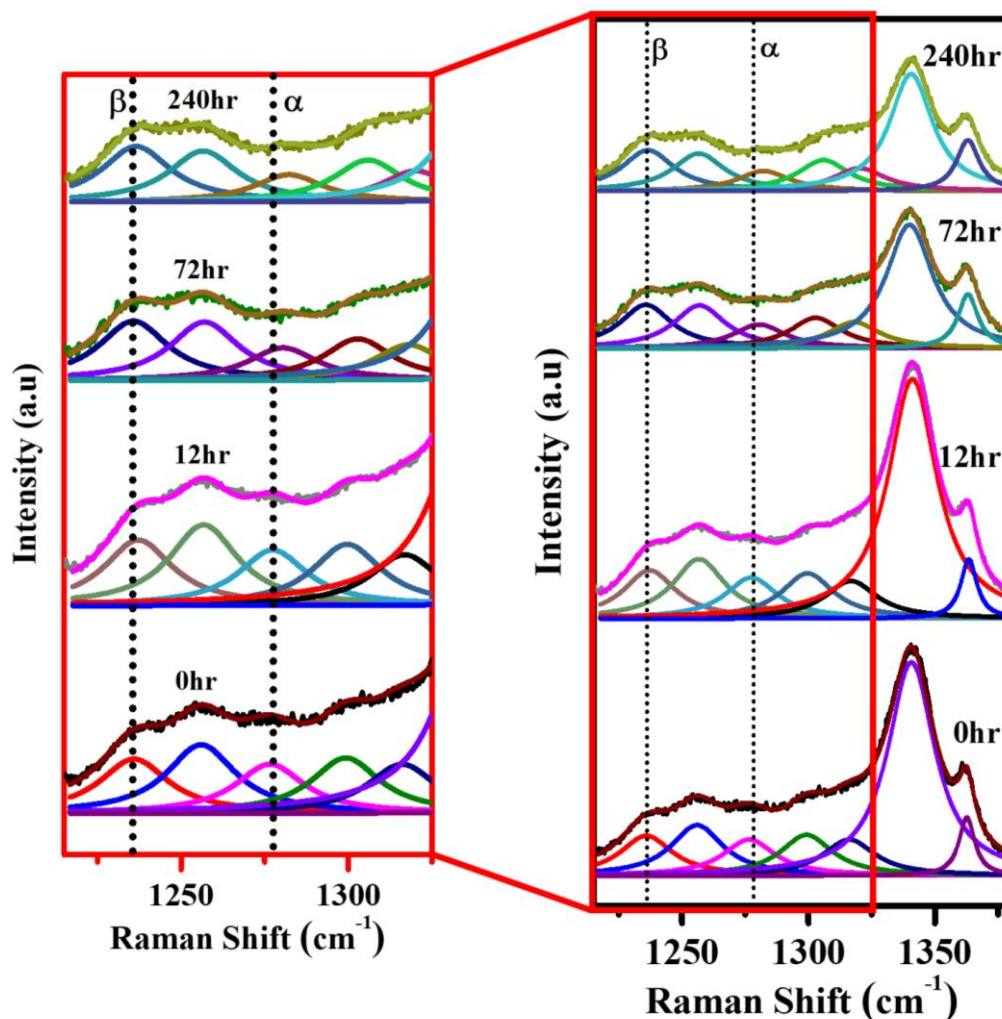


Fig. 3.19. Stack plot with Lorentzian fitting of the Amide III Raman band at selected time points of 0 hr, 12 hr, 72 hr and 240 hr for HEWL incubated in pH 12.2 buffer.

3.3.2.2. Effect of iodoacetamide on the aggregation

Many small molecules like ligands can selectively bind to the native protein, thus preventing significant changes to its original structure when physiochemical changes occur in their environment. In the past, molecules like DTT,⁴³ tri-N-acetylchitotriose,⁴⁴ and iodoacetamide⁴⁵ have demonstrated that HEWL doesn't aggregate in alkaline pH in their presence. We studied the behaviour of 120 μM HEWL

in pH 12.2 in the presence of the small molecule iodoacetamide. As seen in Fig. 3.20a, there is no shift in the amide I band of the protein even on day 10. This confirms that the iodoacetamide molecule is indeed inhibiting protein aggregation. Iodoacetamide is known to attach itself to the dangling $-S-H$ bond.⁴⁶ From this, we learn that the free dangling $-S-H$ bonds of two different monomers were essential for forming new disulphide bonds that led to the polymerisation of the protein.

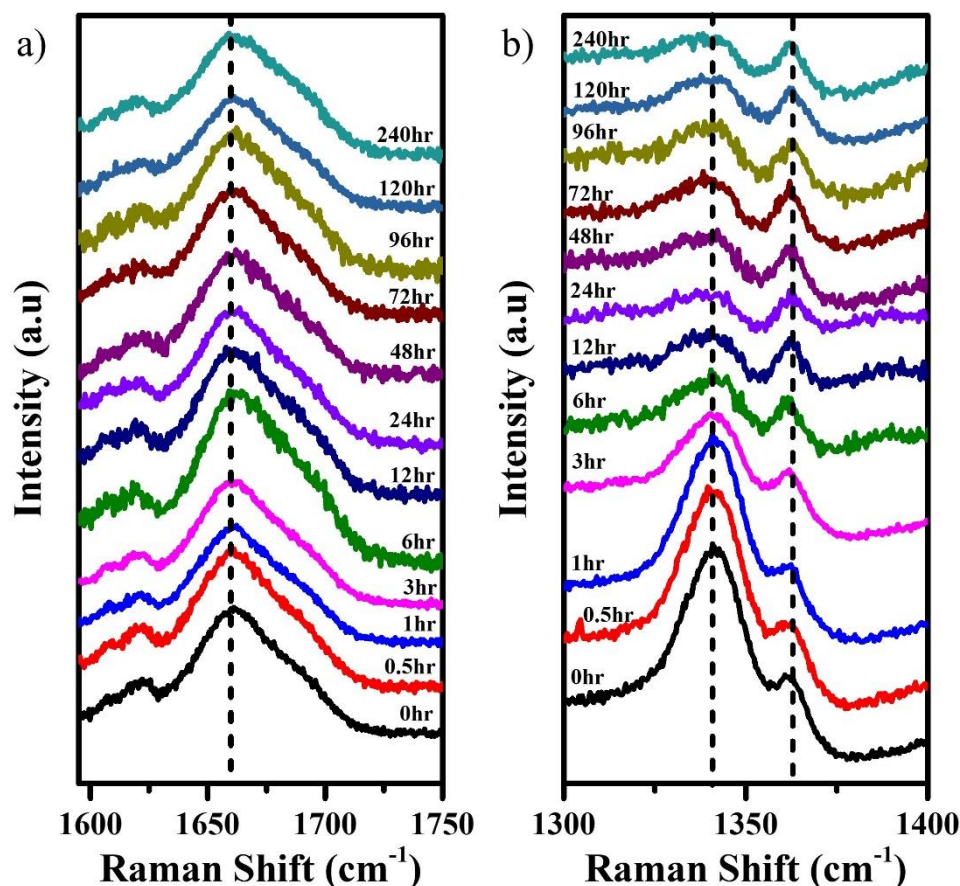


Fig. 3.20. Raman studies of HEWL in pH 12.2 in the presence of iodoacetamide. a) Stack plot of the amide I band, and b) stack plot of the 1341 and 1363 cm^{-1} fermi doublet of the Tryptophan, at various time points of incubation.

3.3.2.3. Other Raman modes analysis

When the 120 μM HEWL protein was incubated in the pH 12.2 buffer, the intensities, shifts in peak positions and the changes in the FWHM of various Raman peaks as a function of time were monitored to reveal more information on the behaviour and environmental effects on their corresponding vibrational modes.

The intensity ratio of the peaks at 1341 cm^{-1} and 1363 cm^{-1} serves as a hydrophobicity marker.^{5, 14, 40} As seen in Fig. 3.21a, for HEWL in pH 7.0, the ratio remained constant and low, confirming that the protein has retained its structure and remains folded. While in Fig. 3.21b, it can be observed that the intensity ratio right from the beginning was high, indicating that the protein had started to unfold, thus exposing the Tryptophans and making them hydrophilic. After 24 hours, the ratio decreases (Fig. 3.21c), indicating that the initially solvent-exposed tryptophans get more buried inside the aggregating protein. In the presence of iodoacetamide, the Tryptophans get buried much quicker (Fig. 3.20b). At time 0 hr, the protein undergoes an alkaline shock causing the disulphide bonds to break, thus exposing the tryptophans to the solvent. When iodoacetamide is first added at 2 hr, the small molecule binds itself to the dangling $-S-H$ bond, preventing any further opening of the protein, thus preserving the conformation of the monomer. The tryptophans get completely buried inside the monomer quickly over the subsequent additions of the iodoacetamide. This result, along with an absence of a shift in the amide I band over 240 hr, indicates that in the presence of iodoacetamide, the HEWL protein prefers to remain monomeric, even in an alkaline pH 12.2 buffer.

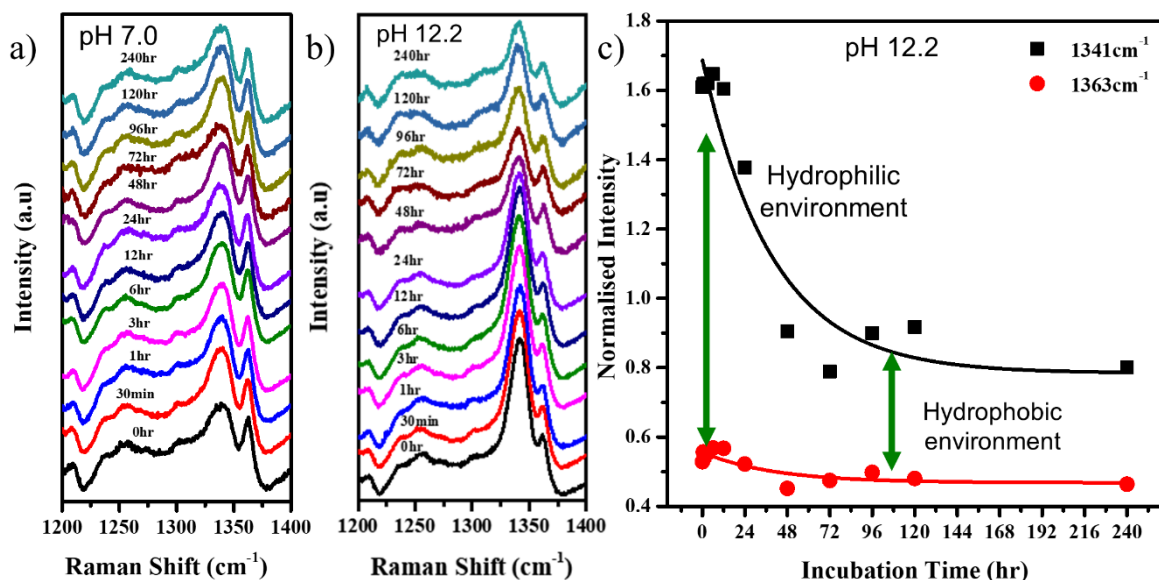


Fig. 3.21. Stack plot of the Raman spectra in the region around $1200 - 1400\text{ cm}^{-1}$, for HEWL in a) pH 7.0 buffer and b) pH 12.2 buffer; c) Time evolution of the intensities of the 1341 and 1363 cm^{-1} Raman peaks corresponding to the Tryptophan residues for the protein in pH 12.2.

As shown in Fig. 3.22a, the FWHM of the 759 cm^{-1} peak originates from the coupled vibrations of in-phase breathing of benzene and pyrrole in the indole ring, increasing from about 9 cm^{-1} to 12 cm^{-1} over as a function of time of aggregation. This indicates a weakening in the coupling due to the distortion of the indole ring.⁴⁰ Hence, Trp conforms in different orientations w.r.t the initial conformations. This is supported by a slight increase in the peak position of the Tryptophan-indole ring's nitrogen around 876 cm^{-1} (Fig. 3.22b).

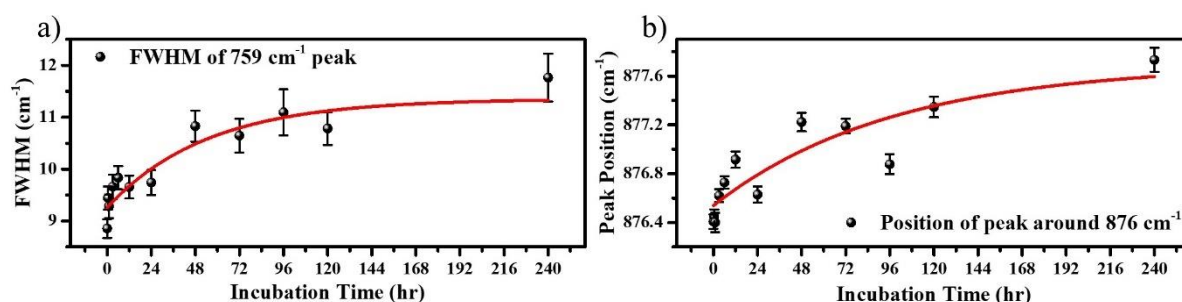


Fig. 3.22. Time evolution of the a) FWHM of the 759 cm^{-1} peak and b) Position of the peak around 876 cm^{-1} .

The 1446 cm^{-1} peak corresponds to the CH, CH_2 , and CH_3 deformation and scissoring mode, and as seen in Fig. 3.23, the FWHM of this peak increases after the initial alkaline shock indicating an opening up of the protein due to increasing deformations.

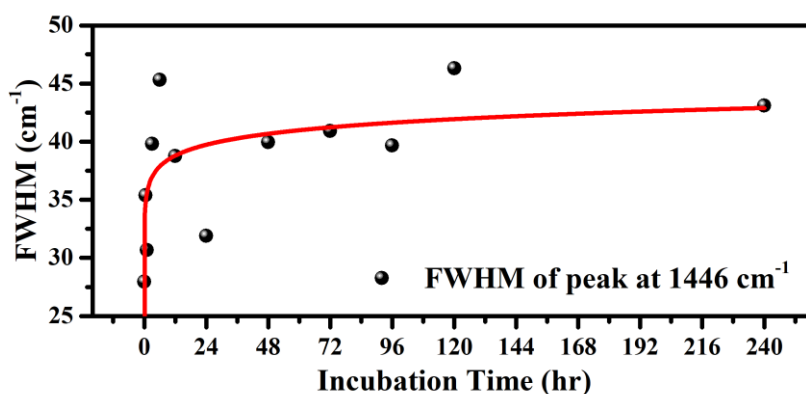


Fig. 3.23. Time evolution of the FWHM of the 1446 cm^{-1} peak.

As seen in Fig. 3.24, the peak around 830 cm^{-1} corresponding to the Fermi doublet associated with the hydroxyl (OH) of tyrosine residues shifts from 834.5 cm^{-1} to 829.7 cm^{-1} , indicating an increased hydrogen bonding around that tyrosine.

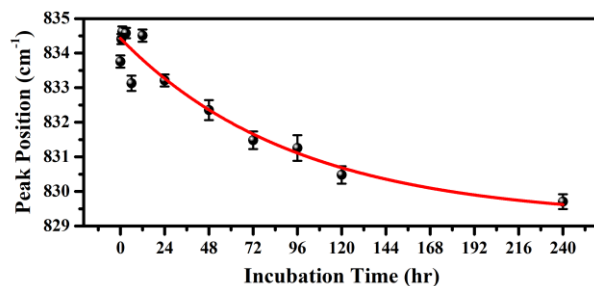


Fig. 3.24. Time evolution of the Raman shift of the peak around 830 cm^{-1} .

The 900 cm^{-1} is a small peak corresponding to the N-C $_{\alpha}$ -C stretching vibration in the skeleton of the alpha-helical part of the secondary structure. As seen in Fig. 3.25a and Fig. 3.25b, an increase in the peak position and its FWHM implies a conformational transformation of the N-C $_{\alpha}$ -C skeleton.

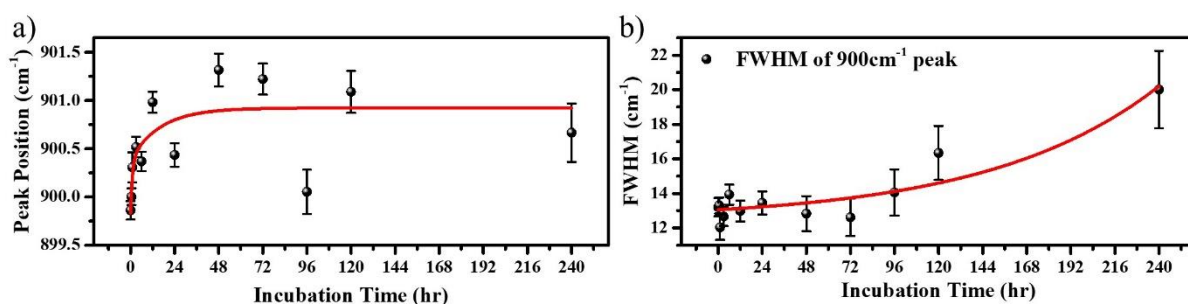


Fig. 3.25. Time evolution of the a) peak position and b) FWHM of the peak around the 900 cm^{-1} Raman peak.

The 932 cm^{-1} peak also corresponds to a similar vibrational mode seen in the skeleton, whose intensity is proportional to the population of the alpha-helical structures.⁵ As seen in Fig. 3.26, the intensity of this peak decreases over time, indicating an overall loss in the alpha-helical content of the protein.⁴⁰

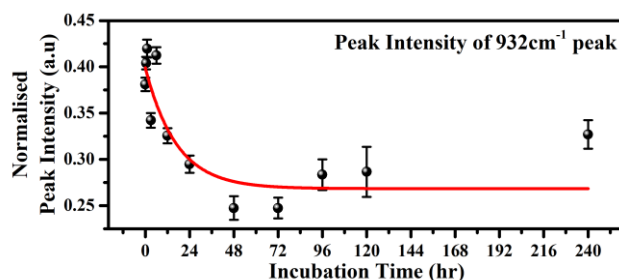


Fig. 3.26. Time evolution of the peak intensity of the 932 cm^{-1} peak.

So far, from the analysis of various Raman modes, we have obtained information on different aspects of the protein in an alkaline environment as a function of time. To better understand the various protein-protein interactions that lead to aggregation and visualising the formed aggregates, we performed MD simulations.

3.3.3. Protein aggregation and interactions – MD simulations[§]

The protein unfolding is followed by the first aggregation step, i.e. dimerization. The protein evolved from the protein unfolding was docked. A dimer with the largest cluster size was selected, and an MD simulation for 200 ns was conducted. The MD simulation trajectory analysis showed that N had a more stable trajectory than HBB from the root mean square deviation graph, as seen in Fig. 3.27c. The aim of increasing the temperature of the simulation was to hasten the unfolding process, as it would be impractical to conduct simulation at room temperature for time scales of a few hours. From these plots in Fig. 3.27, we see that the temperatures 300 K and 380 K are insufficient to unfold the protein, while at 450 K, the protein begins to unfold well within 50 ns. Though N also has a high value of RMSD through the simulation, it is experimentally not feasible hence is ruled out.

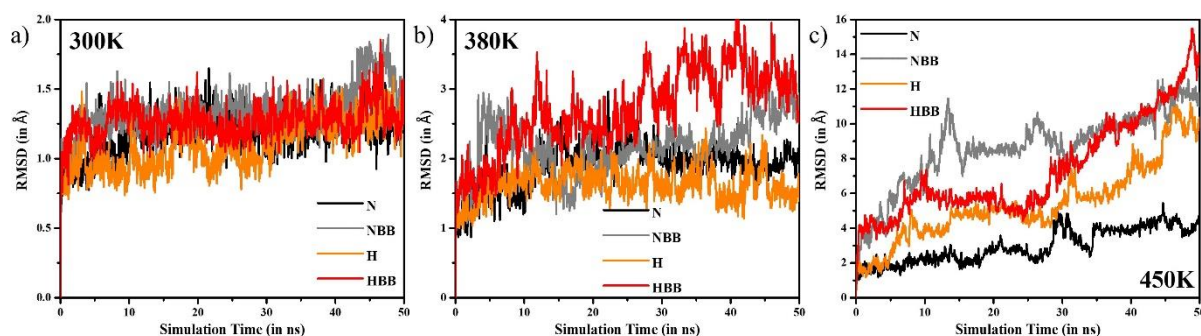


Fig. 3.27. Deviations in the simulation trajectory for the protein in all variations of N, NBB, H and HBB were conducted at a) 300 K, b) 380 K, and c) 450 K.[§]

The last 25 ns of the simulation was very stable in both N and HBB; hence their corresponding frames were considered for interchain interaction analyses and binding free energy calculation for the protein aggregate. The first step to understanding the aggregation starts by analysing the monomer interactions. Non-covalent interactions (hydrogen bond, salt-bridge, pi-cation, pi-pi) plays an essential role in stabilizing the docked structure.⁴⁷⁻⁴⁹ HBB has almost double the interactions as N, proving a favourable association of the monomers in its case (Fig. 3.28a).

[§] Simulations performed by Dr. Shijulal's group from RGC, Thiruvananthapuram

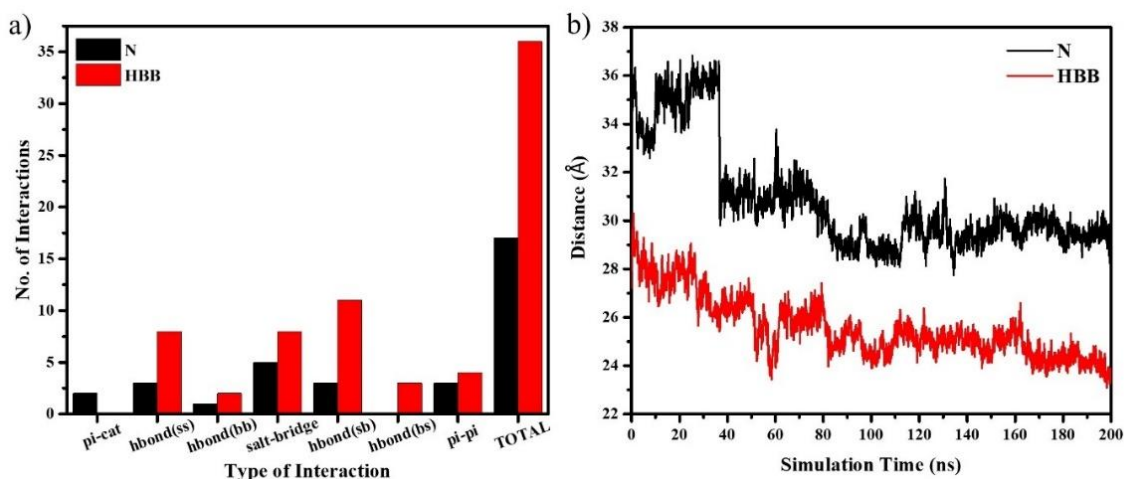


Fig. 3.28. a) The total number of interactions in the dimers, and b) COM as a function of time, in the N and HBB systems.[§]

Analysing the centre of mass (COM) distance w.r.t the simulation time provides insight into the stability of the protein in the dimeric state. From Fig. 3.28b, it is seen that the COM between the monomers of HBB reduced from 27.19 Å to around 23 Å over the simulation duration and stabilized there. A sudden drop from 33.78 Å in the COM distance around 30 ns was observed in the N system but later remained stable around 30 Å. The COM distance does provide an idea of inter-monomeric adjustments to attain stable configuration and supports the fact that an alkaline pH promotes protein aggregation. Binding free energy calculation using the MM-GBSA shows higher affinity among the chains of HBB with the mean binding free energy of -77.53 kCal/mol, whereas the mean binding free energy of N is -62.93 kCal/mol (Fig. 3.29).

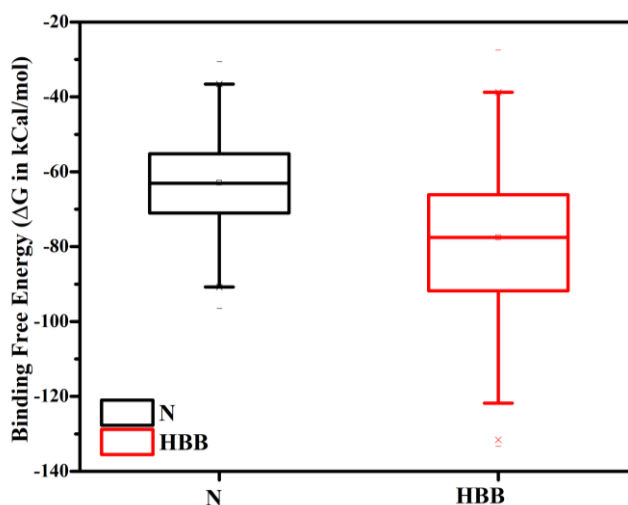


Fig. 3.29. Binding free energy of the docked proteins with the monomers N and HBB.[§]

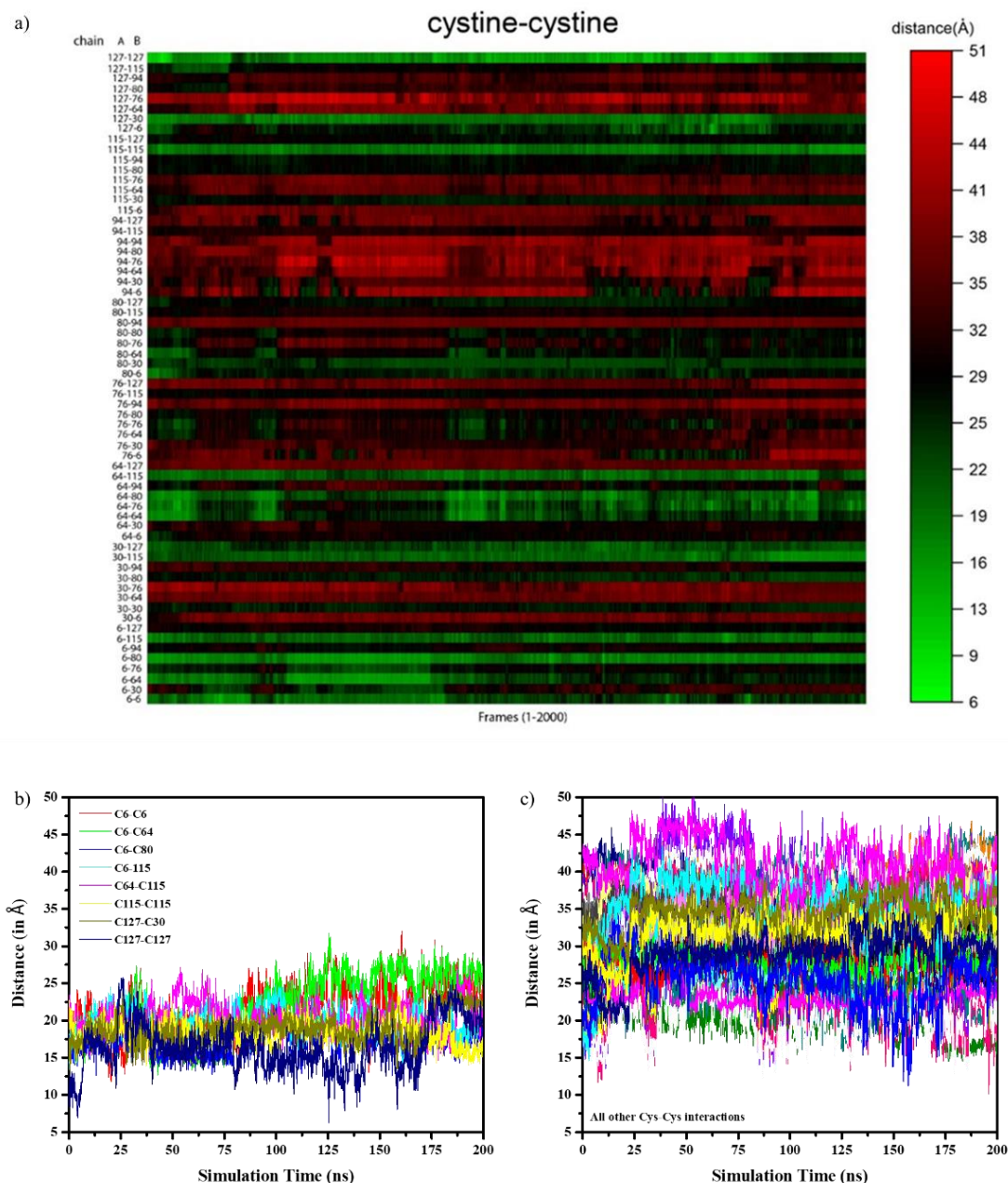


Fig. 3.30. a) Heat map representing all the various cysteine -cysteine distance residues. The distance map from 6 Å (green) representing the closeness of the cysteines from 2 different monomers to 51 Å (red), representing the farther away cysteines, can be used to identify the Cysteine residues that come close. b) The line plots corresponding to the closest occurring cysteine residues in the course of the 200ns of the simulation, c) The line plots corresponding to every other cysteine residue interaction during the simulation.⁸

We mapped the Cys–Cys distances for all the possible 64 combinations over the 200 ns of the simulation (Fig. 3.30a). It was interesting to note that the terminal Cys residues came significantly close (Fig. 3.30b), while the remaining represented in Fig. 3.30c always were at least 10 Å apart. Though they did not come close enough to establish a covalent bond between them, the closeness they developed allowed us to

predict the possibility of their involvement in the disulphide bond formation and hence aggregation.

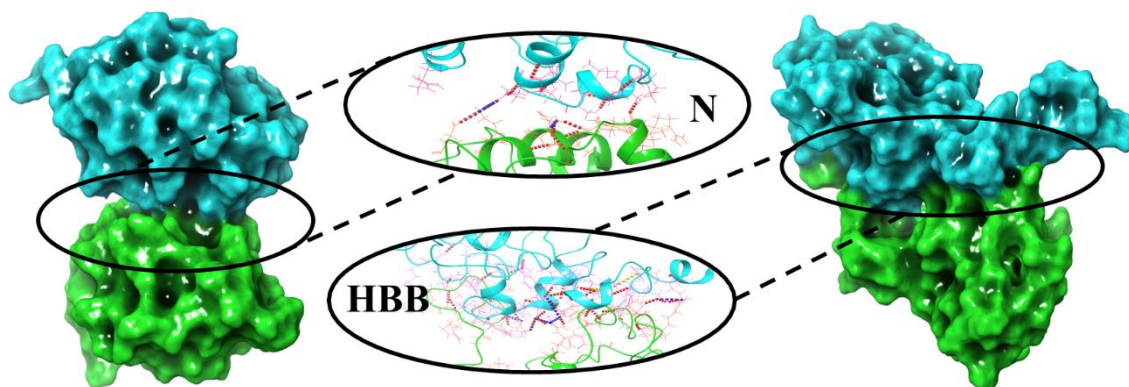


Fig. 3.31. A visual representation of the docked protein dimers in N and HBB indicates a more significant number of interactions in HBB as compared to N.[§]

In Fig. 3.31, we can see a visual of the docked structure after 200 ns of the simulation. Here it is visible that there are many inter-monomeric interactions amongst the monomers in HBB compared to the very few in N. Hence, be assured that the HBB prefers to dimerize in an alkaline environment, while N likes to remain monomeric.

3.4. Aggregation of Human Lysozyme in alkaline pH

Mutations in human lysozyme are associated with systemic non-neuropathic amyloidosis deposits in the kidney, liver, and gastrointestinal tract as amyloid fibrils. Diseases like Alzheimer's, Parkinson's, Huntington's occur due to abnormal protein aggregates in the neural tissues. Since most neurodegenerative diseases occur by amyloid aggregate formation, their diagnosis and treatment depend on the molecular level understanding of the aggregation mechanism.

So far, we have established that Raman spectroscopy can be used to look at some of the molecular level changes occurring in HEWL present in an alkaline pH buffer. Before we study the aggregation behaviour of the HL, we compared the Raman spectra of HEWL and HL to assess their similarities and differences. The experiments were conducted exactly like the HEWL protein, with the incubation in pH 12.2 buffer as the experimental and pH 7.0 as a control system.

3.4.1. Comparing the Raman spectra of HEWL and HL

HEWL and HL are proteins with 129 and 130 amino acids, respectively, as represented in Fig. 3.32. Though they have a 40 % difference in their amino acid composition, they have similar secondary structures, as seen by XRD and NMR.

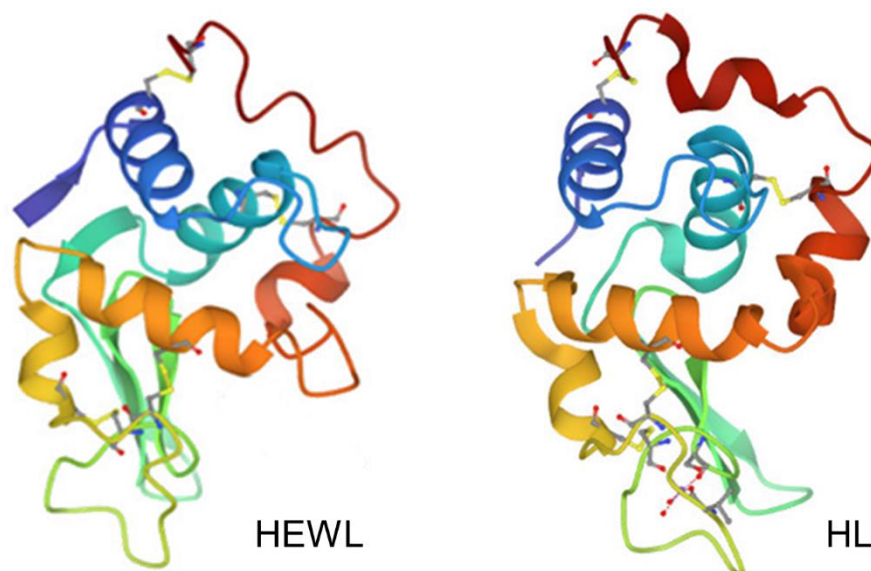


Fig. 3.32. Comparison of PDB structures of HEWL (1HEW) and HL (1LHK).

The Raman spectra of the HEWL vs HL were compared to understand the differences in their native structure that would occur due to their compositional differences, as seen in Fig. 3.33.

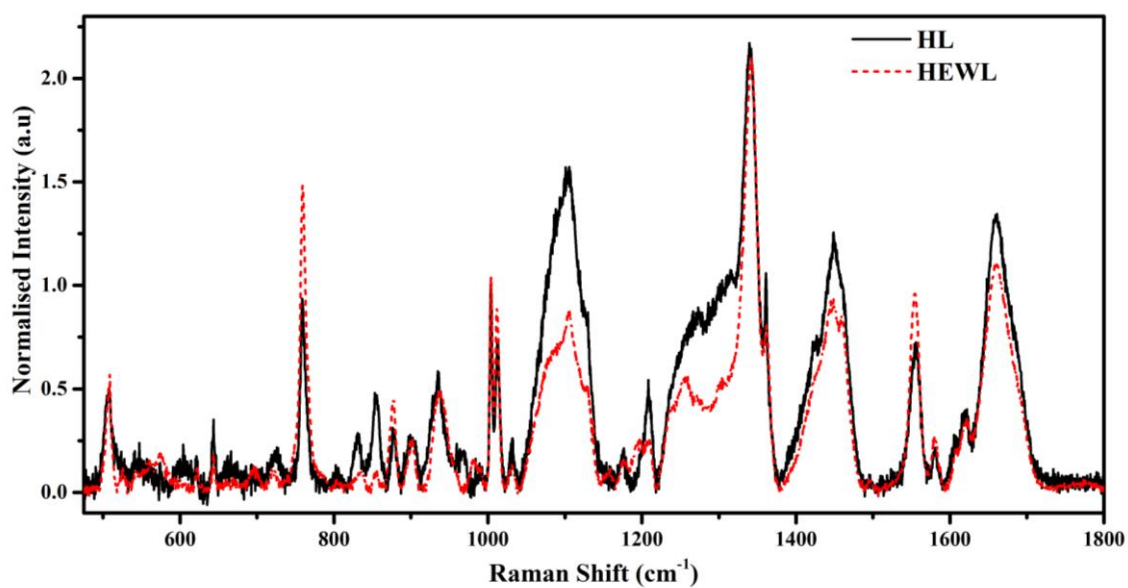


Fig. 3.33. Raman spectra of the HEWL vs. HL at 0 hr incubated in pH 12.2 buffer.

The spectra look very similar but have a few subtle differences on comparing. These differences can be mapped directly to their molecular compositional variations.

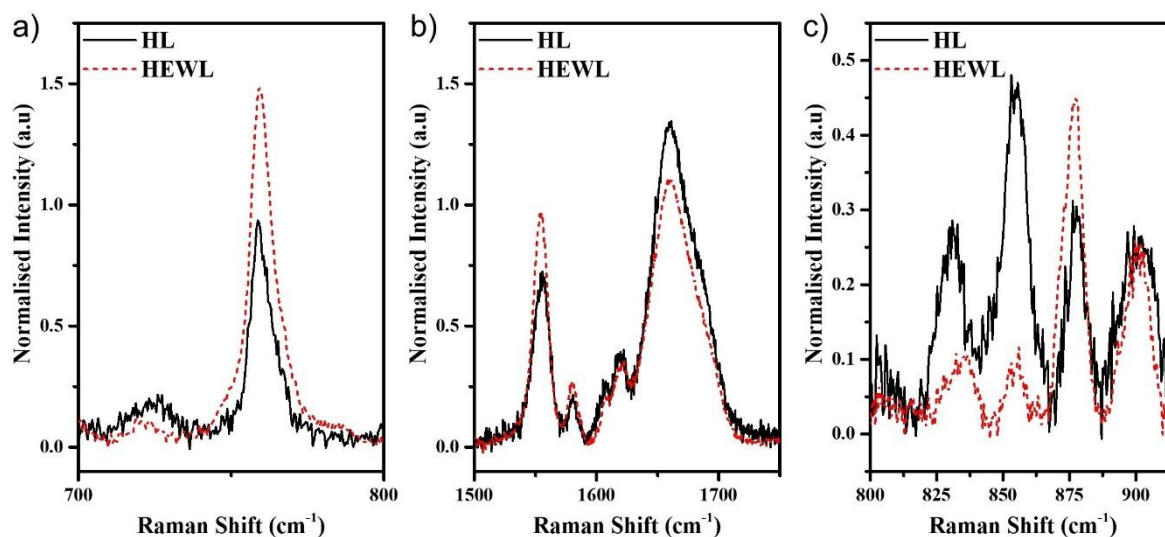


Fig. 3.34. Comparison of the Raman modes corresponding to amino acid side chains and the secondary structure of HEWL and HL proteins, a) Tryptophan indole ring vibrations, b) Tryptophan C=C stretching and the amide I band, and c) Fermi resonance ring breathing modes of Tyrosine.

As seen in Fig. 3.34a and Fig. 3.34b, the intensity of the peak corresponding to tryptophan is higher in HEWL than HL, correctly correlating to 6 tryptophans in HEWL compared to 4 in HL.⁵⁰ The intensity ratio I_{856}/I_{836} is sensitive to the hydrogen bonding and the ionisation state of tyrosine (Fig. 3.34c). This ratio is 1.5 in HL compared to 1.0 in HEWL, indicating that all the 6 tyrosines are exposed to solvent in HL compared to only two out of three in HEWL.⁵⁰

The peaks of the protein's backbone were compared for both HEWL and HL. As seen in Fig. 3.35a, the peak at 900 cm^{-1} corresponds to the N-C α -C stretch in the protein's backbone. The intensity of this peak is a direct measure of the alpha-helical content in the protein. As seen in Fig. 3.35a, the peak intensity is the same for both HEWL and HL, confirming the prediction from XRD and NMR that their secondary structures are similar. A large difference in the intensity of the peak around 1100 cm^{-1} indicates drastic changes in the peptide backbone vibration.⁵⁰ Also, as seen in Fig. 3.35b, the difference in the intensities and shapes of the 1300 cm^{-1} (CH vibration) and 1450 cm^{-1} (CH₂, CH₃ vibration) indicates two different side-chain organisations.⁵⁰

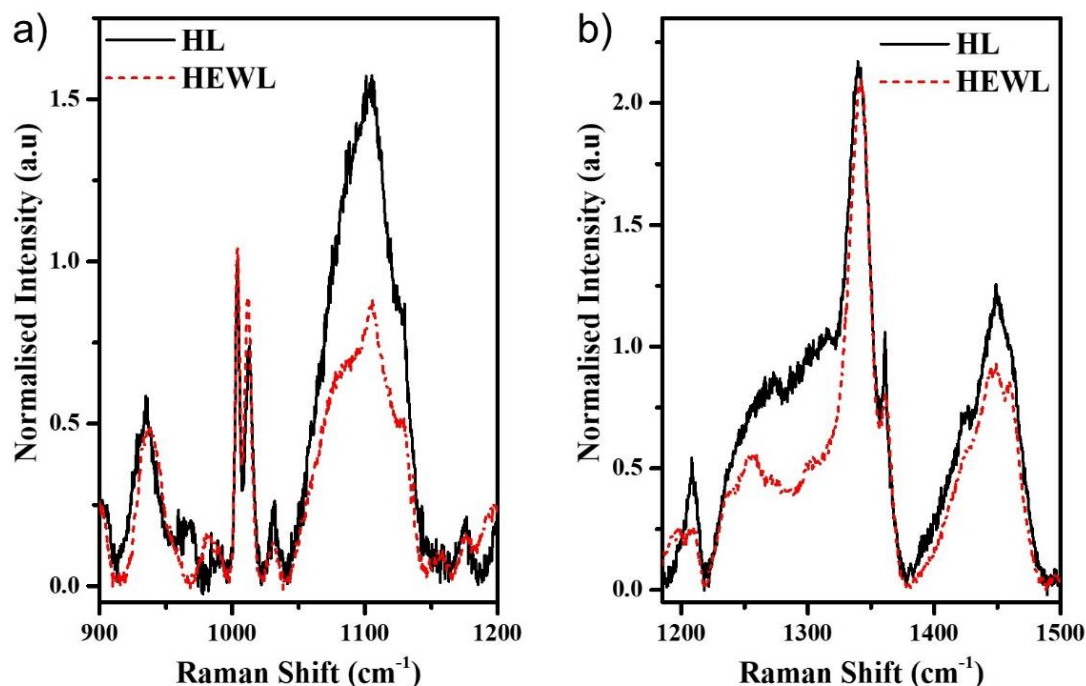


Fig. 3.35. Comparison of the Raman modes corresponding to the backbones of the HEWL and HL proteins, a) N-C α -C stretch and the CH₂ symmetric rock & C α -C stretching, b) CH, CH₂ and CH₃ deformation and scissoring modes.

3.4.2. Amide I band analysis of HL aggregation

As discussed earlier, the amide I band of a protein is used to monitor the secondary structure content using Raman spectroscopy. Like the previous HEWL aggregation studies, 120 μ M HL in pH 12.2 buffer was incubated for 240 hr. The amide I band was monitored at different time intervals. In HEWL, a clear peak shift was observed over time, while in HL, that was not the case. As shown in Fig. 3.36a, there is no shift in the position of the overall amide I band; rather, there is an increasing shoulder around the 1670-1685 cm⁻¹ region. The peak for organised beta-sheets appears in this region. So from the deconvolution of the amide I band into multiple peaks, from about 24 hr, we can see an increase in the beta-sheet content, while there is an overall fall in the alpha-helix and unordered coil content (Fig. 3.36b). Thus from the amide I band analysis, it is evident that the aggregation of the HL occurs similarly to the HEWL by the formation of beta-sheets in the polymers.

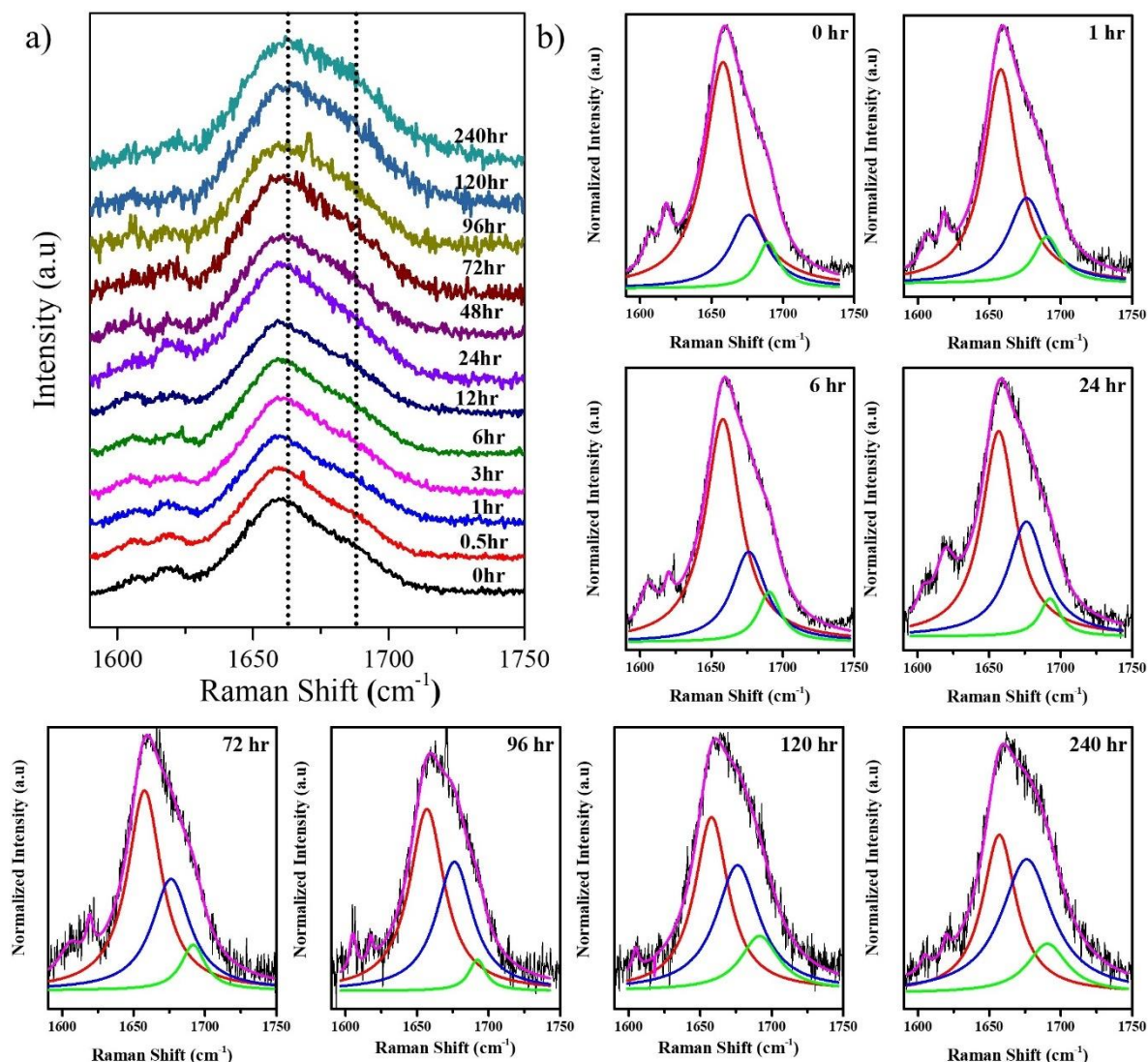


Fig. 3.36. a) Stack plot of the amide I band of HL incubated in an alkaline buffer at pH 12.2, b) deconvolution of the amide I band into various components at different time points, red – alpha-helix and unordered structure, blue – beta-sheet content, green – beta turns.

3.5. Conclusion

In our present study, through Raman spectroscopic investigations and Molecular Dynamic simulations, we were able to gain insight into the various behavioural changes of the protein's backbone, the amino acids, and the other inter-/intra- molecular bonds. In brief, the environment of the Tryptophan residues analysed from the I_{1341}/I_{1363} gives us insight into the conformation of the protein, i.e. whether it is unfolded or folded. Analysing various backbone peaks like the 900 cm^{-1} , 932 cm^{-1} , 1446 cm^{-1} show us that the system is going into multiple conformations and has a decreasing alpha-helical content. The peaks of the amino acid residues like 759 cm^{-1} ,

876 cm^{-1} and 832 cm^{-1} demonstrate the environmental changes of these residues, thus pointing to their positions during and after aggregation. Overall, by observing the changes in various Raman modes, we can understand the behaviour of different components, including the backbone, the side chains and different intra-/inter-molecular bonds. Thus, giving us an insight into the various behaviours during the aggregation mechanism and after it. These multiple views prove how versatile a single tool like Raman spectroscopy can be in understanding protein aggregation, compared to the many other tools that provide few insights. MD simulations give us insight into the molecular details of the model system subjected to a series of changes. Hence, act as a reliable tool to predict molecular activities responsible for protein aggregation. We used various modes of analysis to predict changes in the HEWL system that promote aggregation. The secondary structure analysis of the protein monomer at 450K helps establish that the protein in pH 12.2 unfolds the most due to the de-stability of intramolecular bonds. The partially unfolded structures have exposed amino acids that allow for inter-monomeric interactions. This is reflected in the non-bonded interactions map, which shows almost double the number of interactions in HBB than in N. The inter-monomeric centre of mass distance plot shows that the HBB monomers come closer over the simulation time, indicating a favourable dimer formation. The choice of broken disulphide bonds in a pH 12.2 environment supports the aggregation mechanism, proven in the high dimer affinity in the docked protein structures. Overall, we observed that the protein tends to aggregate in an alkaline pH environment at room temperature, from the subtle and distinctive changes in the Raman spectra over the ten days of the experiment and the protein docking studies. After establishing this, we used Raman spectra to understand molecular-level differences between the human lysozyme and the HEWL. The time-dependent studies of HL in pH 12.2 revealed that it also aggregates by forming beta-sheets.

3.6. References

1. Fox, L.M. & Yamamoto, A. in *Autophagy: Cancer, Other Pathologies, Inflammation, Immunity, Infection, and Aging* (ed. Hayat, M.A.) 117 (Academic Press, Amsterdam, 2015).
2. Philo, J. & Arakawa, T. Mechanisms of Protein Aggregation. *Current pharmaceutical biotechnology* 10, 348 (2009).
3. Merlini, G., Bellotti, V., Andreola, A., Palladini, G., Obici, L., Casarini, S., Perfetti, V. Protein Aggregation. *Clinical Chemistry and Laboratory Medicine* 39, 1065 (2001).
4. Ettah, I. & Ashton, L. Engaging with Raman Spectroscopy to Investigate Antibody Aggregation. *Antibodies (Basel, Switzerland)* 7, 24 (2018).
5. Dolui, S. et al. Order, Disorder, and Reorder State of Lysozyme: Aggregation Mechanism by Raman Spectroscopy. *The Journal of Physical Chemistry B* 124, 50 (2020).
6. Kurouski, D., Van Duyne, R.P. & Lednev, I.K. Exploring the structure and formation mechanism of amyloid fibrils by Raman spectroscopy: a review. *Analyst* 140, 4967 (2015).
7. Dumetz, A.C., Chockla, A.M., Kaler, E.W. & Lenhoff, A.M. Effects of pH on protein-protein interactions and implications for protein phase behavior. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* 1784, 600 (2008).
8. Ravi, V.K., Swain, T., Chandra, N. & Swaminathan, R. On the Characterization of Intermediates in the Isodesmic Aggregation Pathway of Hen Lysozyme at Alkaline pH. *PLOS ONE* 9, e87256 (2014).
9. Lewis, E.N. et al. Combined Dynamic Light Scattering and Raman Spectroscopy Approach for Characterizing the Aggregation of Therapeutic Proteins. *Molecules* 19 (2014).
10. Bagriantsev, S., Kushnirov, V. & Liebman, S. Analysis of Amyloid Aggregates Using Agarose Gel Electrophoresis. *Methods in enzymology* 412, 33 (2006).
11. Adawy, A. & Groves, M.R. The Use of Size Exclusion Chromatography to Monitor Protein Self-Assembly. *Crystals* 7 (2017).
12. Fekete, S., Beck, A., Veuthey, J.-L. & Guillaume, D. Theory and practice of size exclusion chromatography for the analysis of protein aggregates. *Journal of Pharmaceutical and Biomedical Analysis* 101, 161 (2014).

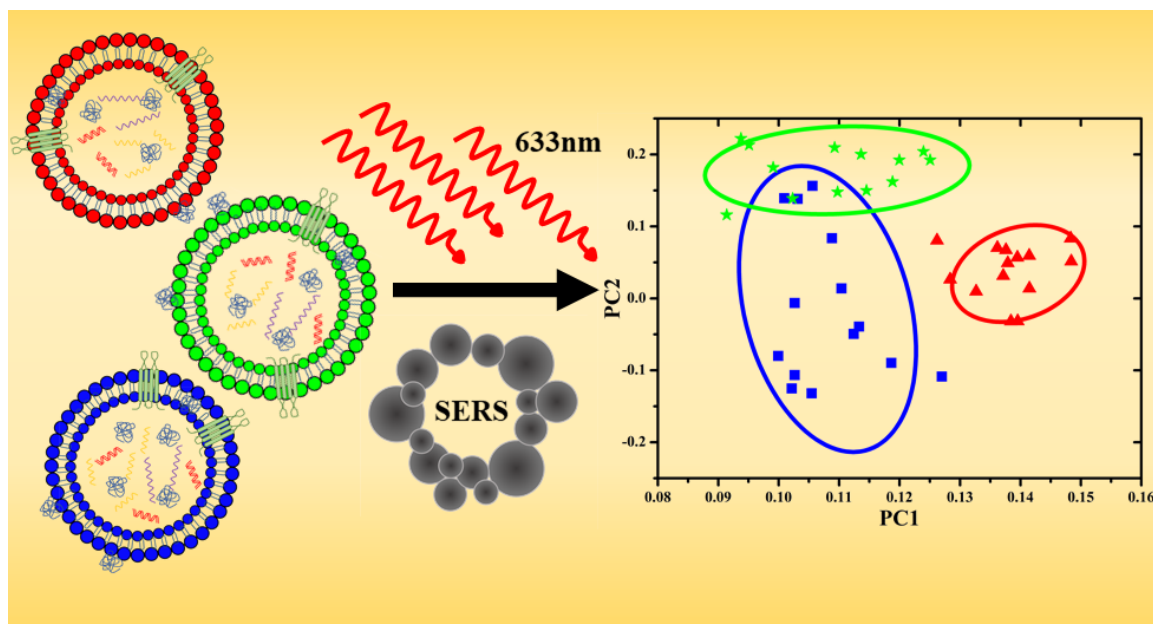
13. Hong, P., Koza, S. & Bouvier, E.S.P. Size-Exclusion Chromatography for the Analysis of Protein Biotherapeutics and their Aggregates. *Journal of liquid chromatography & related technologies* 35, 2923 (2012).
14. Xing, L. et al. Double-edged effects of aluminium ions on amyloid fibrillation of hen egg-white lysozyme. *International Journal of Biological Macromolecules* 132, 929 (2019).
15. Ebrahim-Habibi, A., Morshedi, D., Rezaei-Ghaleh, N., Sabbaghian, M. & Nemat-Gorgani, M. Protein-Protein interactions leading to aggregation: Perspectives on mechanism, significance and control. *Journal of the Iranian Chemical Society* 7, 521 (2010).
16. Arnaudov, L.N. & de Vries, R. Thermally Induced Fibrillar Aggregation of Hen Egg White Lysozyme. *Biophysical Journal* 88, 515 (2005).
17. Karamanos, T.K., Kalverda, A.P., Thompson, G.S. & Radford, S.E. Mechanisms of amyloid formation revealed by solution NMR. *Progress in nuclear magnetic resonance spectroscopy* 88-89, 86 (2015).
18. Bronsoms, S., Trejo, S.A. in *Insoluble Proteins. Methods in Molecular Biology (Methods and Protocols)* (ed. García-Fruitós, E.) (Humana Press, New York, NY, 2015).
19. Wani, A.H. & Udgaonkar, J.B. Mass spectrometry studies of protein folding. *Current Science* 102, 245 (2012).
20. Chaari, A., Fahy, C., Chevillot-Biraud, A. & Rholam, M. Insights into Kinetics of Agitation-Induced Aggregation of Hen Lysozyme under Heat and Acidic Conditions from Various Spectroscopic Methods. *PLOS ONE* 10, e0142095 (2015).
21. Eugenia, P. et al. Raman Spectroscopy of Protein Crystal Nucleation and Growth. *American Journal of Biochemistry and Biotechnology* 10, 202 (2014).
22. Niaura, G. Raman Spectroscopy in Analysis of Biomolecules. *Encyclopedia of Analytical Chemistry*, 1 (2014).
23. Evans, K. Poking Holes in Antibiotic Resistance. (2015).
24. Foggia, M., Taddei, P., Torreggiani, A., Dettin, M. & Tinti, A. Self-assembling peptides for biomedical applications: IR and Raman spectroscopies for the study of secondary structure. *Proteomics Research Journal* 2, 231 (2012).
25. E, L. & A, K. in *Molecular Modeling of Proteins. Methods in molecular biology (Methods and Protocols)* 3 (Humana Press / Springer, 2015).

26. Piana, S., Lindorff-Larsen, K. & Shaw, D.E. Atomic-level description of ubiquitin folding. *Proceedings of the National Academy of Sciences* 110, 5915 (2013).
27. Prigozhin, M.B. & Gruebele, M. Microsecond folding experiments and simulations: a match is made. *Physical Chemistry Chemical Physics* 15, 3372 (2013).
28. Sultan, M.M., Denny, R.A., Unwalla, R., Lovering, F. & Pande, V.S. Millisecond dynamics of BTK reveal kinome-wide conformational plasticity within the apo kinase domain. *Scientific Reports* 7, 15604 (2017).
29. Shaw, D.E. et al. in Proceedings of the Conference on High Performance Computing Networking, Storage and Analysis Article 65 (Association for Computing Machinery, Portland, Oregon, 2009).
30. Redler, R.L. et al. Computational approaches to understanding protein aggregation in neurodegeneration. *Journal of Molecular Cell Biology* 6, 104 (2014).
31. Karplus, M. & McCammon, J.A. Molecular dynamics simulations of biomolecules. *Nature Structural Biology* 9, 646 (2002).
32. Filik, J. & Stone, N. Drop coating deposition Raman spectroscopy of protein mixtures. *Analyst* 132, 544 (2007).
33. Janek, T., Rodrigues, L.R., Gudiña, E.J. & Burger, J. Synergistic effect of hen egg white lysozyme and lysosomotropic surfactants on cell viability and membrane permeability. *Colloids and Surfaces B: Biointerfaces* 185, 110598 (2020).
34. Mach, H., Middaugh, C.R. & Denslow, N. Determining the Identity and Purity of Recombinant Proteins by UV Absorption Spectroscopy. *Current Protocols in Protein Science* 1, 7.2.1 (1995).
35. Zhang, D. et al. Raman Detection of Proteomic Analytes. *Analytical Chemistry* 75, 5703 (2003).
36. Teizo Kitagawa, S.H. in Handbook of Vibrational Spectroscopy (ed. Peter Griffiths, J.M.C.) (John Wiley & Sons, Ltd, 2006).
37. Talari, A.C.S., Movasaghi, Z., Rehman, S. & Rehman, I.u. Raman Spectroscopy of Biological Tissues. *Applied Spectroscopy Reviews* 50, 46 (2015).
38. Rygula, A. et al. Raman spectroscopy of proteins: a review. *Journal of Raman Spectroscopy* 44, 1061 (2013).

39. Wang, C.-H., Huang, C.-C., Lin, L.-L. & Chen, W. The effect of disulfide bonds on protein folding, unfolding, and misfolding investigated by FT-Raman spectroscopy. *Journal of Raman Spectroscopy* 47, 940 (2016).
40. Xing, L. et al. Amyloid formation kinetics of hen egg white lysozyme under heat and acidic conditions revealed by Raman spectroscopy. *Journal of Raman Spectroscopy* 50, 629 (2019).
41. Florence, T.M. Degradation of protein disulphide bonds in dilute alkali. *Biochemical Journal* 189, 507 (1980).
42. Uversky, N.V. Amyloidogenesis of Natively Unfolded Proteins. *Current Alzheimer Research* 5, 260 (2008).
43. Kumar, S., Ravi, Vijay K. & Swaminathan, R. How do surfactants and DTT affect the size, dynamics, activity and growth of soluble lysozyme aggregates? *Biochemical Journal* 415, 275 (2008).
44. Kumar, S., Ravi, V.K. & Swaminathan, R. Suppression of lysozyme aggregation at alkaline pH by tri-N-acetylchitotriose. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* 1794, 913 (2009).
45. Ravi, V.K., Goel, M., Kotamarthi, H.C., Ainaravapu, S.R.K. & Swaminathan, R. Preventing Disulfide Bond Formation Weakens Non-Covalent Forces among Lysozyme Aggregates. *PLOS ONE* 9, e87012 (2014).
46. Wright, M.H. & Sieber, S.A. Chemical proteomics approaches for identifying the cellular targets of natural products. *Natural Product Reports* 33, 681 (2016).
47. Halling, P.J. Proteins: Structures and molecular properties (2nd edition). by Thomas E. Creighton, W. H. Freeman, New York, 1992, xiii + 512 pp, price £22.95. ISBN 0-7167-7030-X. *Journal of Chemical Technology & Biotechnology* 62, 105 (1995).
48. Velev, O.D., Kaler, E.W. & Lenhoff, A.M. Protein Interactions in Solution Characterized by Light and Neutron Scattering: Comparison of Lysozyme and Chymotrypsinogen. *Biophysical Journal* 75, 2682 (1998).
49. Verheul, M., Roefs, S.P.F.M. & de Kruif, K.G. Kinetics of Heat-Induced Aggregation of β -Lactoglobulin. *Journal of Agricultural and Food Chemistry* 46, 896 (1998).
50. Marx, J., Jacquot, J., Berjot, M., Puchelle, E. & Alix, A.J.P. Characterization and conformational analysis by Raman spectroscopy of human airway lysozyme. *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology* 870, 488 (1986).

Chapter 4. Using Surface-Enhanced Raman Spectroscopy to Distinguish Extracellular Vesicles under Autophagic Conditions^{**††}

Extracellular vesicles (EVs) laden with lipid, protein, DNA, micro-RNAs, play important biological functions in intercellular communication and have pivotal roles in pathophysiological conditions. Characterisation of the EVs has always been a multi-step process involving large volumes as they are heterogeneous in size and properties. Here, simple citrate-reduced silver nanoparticle-assisted SERS distinguishes EVs extracted from several cell lines isolated under autophagic conditions (nitrogen starvation). We used two cancerous cell lines, HeLa, its corresponding autophagy-deficient cell line (Atg5^{-/-}), and a non-cancerous cell line, HEK293, to isolate EVs. Our study helps the facile detection and differentiation of EVs isolated between two closely related human cell lines that differ by their autophagic ability. The Principal Component Analysis (PCA) of the SERS spectra of these EVs consistently showed the presence of distinct chemical compositions of the EVs. SERS of EVs can help probe more into the molecular level information from EVs and could become a powerful tool once coupled with improved microscopy techniques for diagnosis and therapy.



^{**} This work was carried out in collaboration with Dr. Sreedevi, and Prof. Rai Manjithaya, MBGU, JNCASR

^{††} This work has been published as “Surface-Enhanced Raman Spectroscopy as a Tool for Distinguishing Extracellular Vesicles under Autophagic Conditions: A Marker for Disease Diagnostics”, Divya Chalapathi *et al*, J. Phys. Chem. B, 2020, 124 (48), 10952-10960. Copyright © American Chemical Society.

4.1. Introduction

Intercellular communication between cells is essential in many physiological roles, from lower prokaryotic organisms such as bacteria to higher eukaryotes. This is brought about by signalling molecules, proteins, and sometimes by EVs too. These EVs are made up of lipids, proteins, DNA, RNA, microRNA, and non-coding RNA. They differ in various aspects and are classified based on their protein markers, shape, size, and origin/biogenesis, including exosomes, micro-vesicles, apoptotic EVs, autophagic EVs (Fig. 4.1).¹⁻⁴ The heterogeneity of these vesicles is due to their subtle differences in their composition. Exosomes are vesicles ranging between 20 and 150 nm secreted from cells and are involved in intercellular communication.^{5, 6} Exosomes retain constituents of a donor cell, including proteins, DNA, miRNA and others, thus preserving the genetic information. Hence they are potential biomarkers in many pathophysiological conditions like cancer and neurodegenerative diseases.^{7, 8}

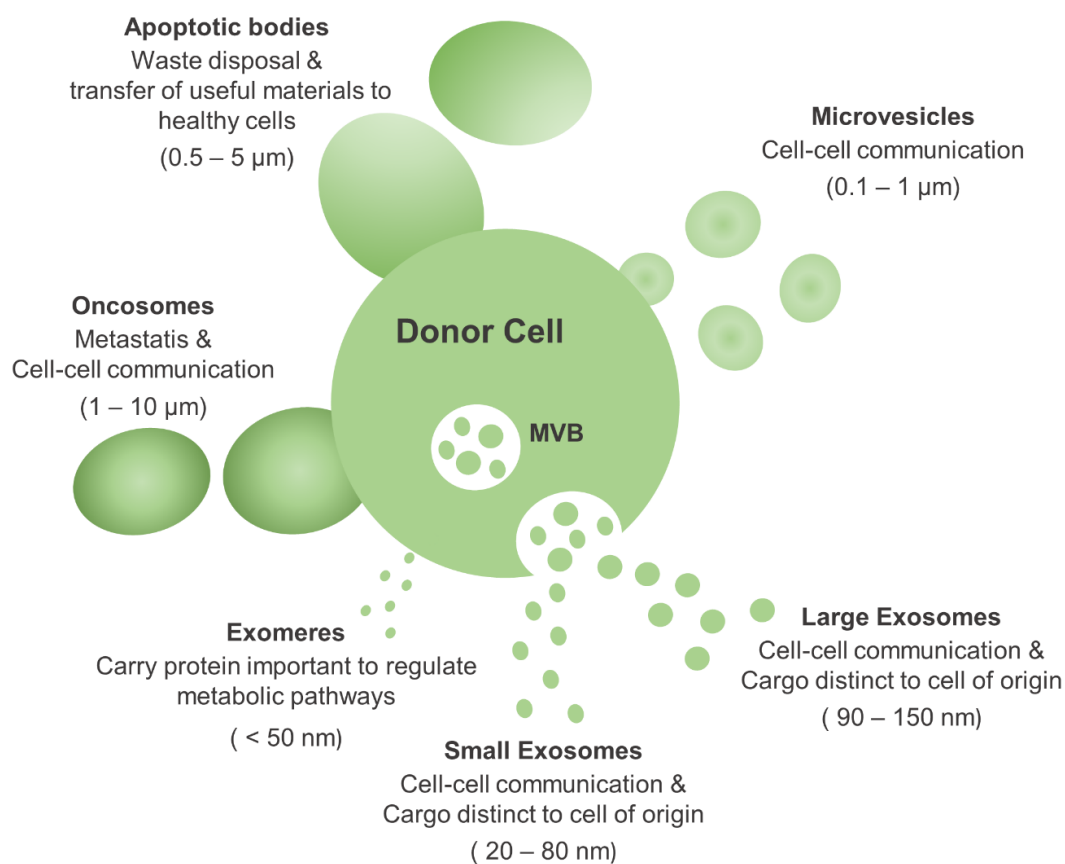


Fig. 4.1. Schematic representation of different EV subpopulations. (Inspired from Zhang, Y. et al ⁹)

Inside a donor cell, the formed vesicles are compartmentalized into the multivesicular bodies, which fuse with the cell membrane and get released into the extracellular space (Fig. 4.2).¹⁰ Recently, it has been suggested that the vesicles carrying all the tetraspanin markers (CD63, CD9, and CD81) shall be termed exosomes and others as just extracellular vesicles (EVs).¹

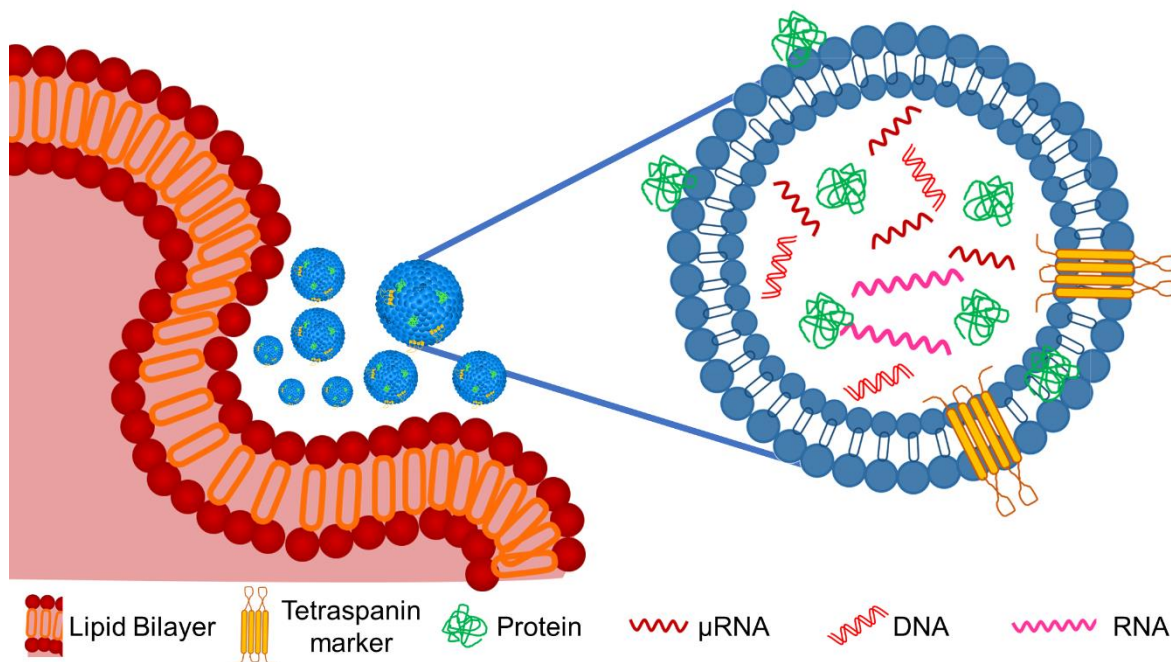


Fig. 4.2. Schematic representation of the release of EVs from the donor cell. The various constituents of the EV are indicated in the image.

Various stress stimuli, like metabolic stress, can activate cellular mechanisms for the cells to either tolerate these adverse conditions or trigger cell suicide mechanisms to eliminate damaged and potentially dangerous cells. These stress stimuli can include nutrition deficiency in starvation conditions, cell death, hyperthermia, and hypoxia. These stresses stimulate a conserved pathway called autophagy, as seen in Fig. 4.3. The formed double-membrane vesicles called autophagosomes engulf proteins, cytoplasm, protein aggregates, and organelles. They are then transported to lysosomes, where they fuse to result in an autolysosome. Here their contents are degraded and recycled, thus providing energy. It plays an essential role in fighting many diseases, including viral and bacterial infections, neurodegenerative disorders, cancer, and cardiovascular diseases.¹¹

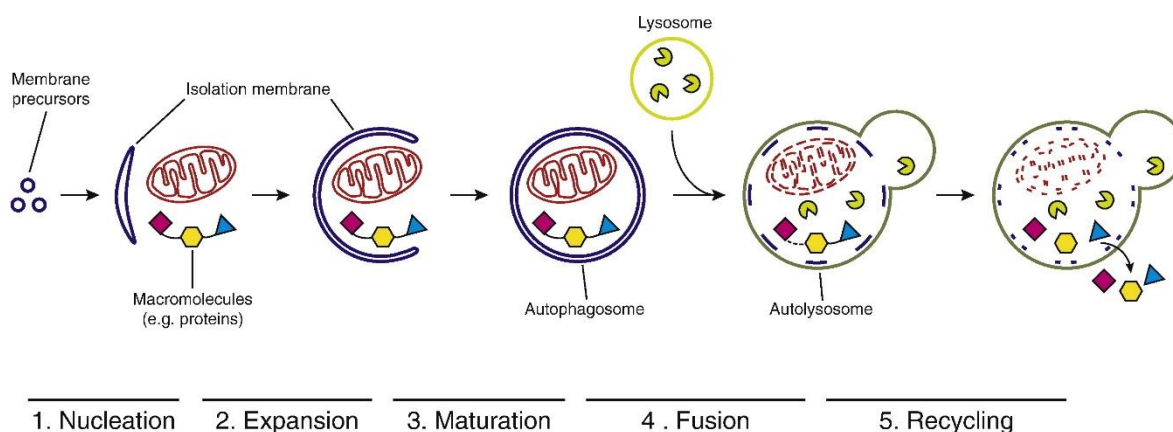


Fig. 4.3. Schematic representation of the selective autophagy mechanism. © 2016, Gabriele Zaffagnini, Sascha Martens. Published by Elsevier Ltd.

Cancer cells are known to be aggressive and survive by resisting apoptotic cell death.¹² Studies have shown that autophagy can be either tumour-suppressing or tumour-promoting.¹³ When recycling pathways are compromised, exosome secretion is an alternative way to alleviate stress, thus inducing autophagy. Reports show that glucose starvation enhances exosome secretion in cardiomyocytes and promotes angiogenesis in endothelial cells.¹⁴ With this background, we wanted to characterise EVs isolated from cancerous cell types grown under autophagic (starvation) conditions that have not been studied so far. Hence, we chose a cancerous cell line HeLa and its corresponding autophagy knockout ($Atg5^{-/-}$). We chose a non-cancerous human embryonic kidney cell line HEK293 as a control system.

The genes containing the autophagy machinery are called ATG (AuTophagy related gene). Various proteins generated by these genes play significant roles in the different steps of the autophagy process. The $Atg5$ is one of the pivotal proteins in the orchestrated cascade of the autophagy process. Though the process of canonical autophagy is primarily degradative, in contrast, secretory autophagy is another manifestation of this process that facilitates the secretion of cytosolic constituents, including proteins, lipids, and nucleic acids. It is reported that the autophagy-related proteins such as $Atg5$ directly regulate the fate of MVBs (Multi Vesicular Bodies) and subsequent exosome biogenesis.¹⁵ Specifically, the knockout of $ATG5/ATG16L1$ but not $ATG7$ has been found significantly to reduce exosome release.¹⁶

Exosomes can be used for personal diagnosis and therapy, and of late, many different techniques have been employed for their detection. These include electron microscopy (TEM and SEM),^{17, 18} Surface Enhanced Raman Microscopy,¹⁹⁻²¹ Surface Plasmon Resonance,²² Fluorescence Live-cell Microscopy, Enzyme-Linked Immunosorbent Assay,²³ Western Blot,²³ Bioanalyzer, Atomic Force Microscope,²⁴ Peak Force Microscopy, Dynamic Light Scattering,²⁵ Nanoparticle Tracking Analysis (NTA),²⁶ Flow Cytometry.²⁷ The high-end techniques like NTA, TRPS (Tunable Resistive Pulse Sensing),²⁸ TIRFM (Total Internal Reflection Fluorescence Microscopy),²⁹ and TEM/SEM help track, visualize, and characterise the physical properties of the extracellular vesicles that get secreted out. Hence, vibrational spectroscopy becomes one of the powerful tools that can be used to characterise EVs at the molecular level.

SERS technique has been successfully employed earlier in detecting various exosomes, the majority of them being extracted from multiple kinds of cancer cells.³⁰ Some studies involve extensive preparation of plasmonic nanostructures for improved SERS enhancement, while others include a capture-detector probe assay for selective SERS studies. Most of these techniques used in earlier studies either require an elaborate assay to prepare the SERS substrate or a capture-detection strategy or extensive labelling to detect the exosomes with specific markers. As the molecular composition of the EVs cannot be distinguished vividly, analysing very closely related Raman spectra provide several challenges for making meaningful interpretations. Principal Component Analysis (PCA) is a technique used to sort such Raman data to obtain differences and hence, get meaningful patterns amongst vast quantities of data.^{19, 31} This chapter reports a facile detection and differentiation of EVs, using SERS spectra obtained by citrate reduced silver nanoparticles that were further analysed by PCA. The data analysis allowed us to get an insight into the chemical compositional variation of the EVs. This kind of resolution obtained can also reveal the molecular details of the EVs, making it a potent tool when combined with an improved microscopic tool for diagnostic and therapeutic applications.

4.2. Materials and Methods

a) Isolation of EVs from the cell culture media:

The 6-well plates containing confluent cells (HeLa, Atg5^{-/-}, HEK293) were subjected to starvation in EBSS (Earle's Balanced Salt Solution) (1 mL) for 2 hours. Atg5^{-/-}HeLa cells were a kind gift from Prof. Richard Youle, NIH. The EVs were isolated from the culture supernatant using TEIR (Total Exosome Isolation Reagent) in the total exosome isolation kit as per the manufacturer's protocol (Invitrogen, Catalog #4478359). The pellets were resuspended in 50 μ l sterile PBS. The cell lysates and EVs isolated from HeLa, Atg5^{-/-}, and HEK293 cells were probed with anti-rabbit ATG5, anti-rabbit GM130 (both Cell Signalling Technology), and anti-mouse CD63 (DSHB). All the primary antibodies were used in 1:1000 dilution, and the secondary antibodies were used in 1:10000 dilution. GM130 was used as a control for the isolation procedure in the western blot.^{‡‡}

b) SERS – Sample preparation and spectroscopy parameters:

The AgNPs were centrifuged at 7200 rpm for 10 min to concentrate 10X. A 10 μ l mixture of one part of EVs with nine parts of concentrated AgNPs was incubated on ice for 8-10 minutes. The incubated mixture was dropped onto a siliconized glass substrate and placed on the sample stage. The drop was randomly scanned using a 20X long working distance objective to focus the laser beam to an approximate diameter of 1 μ m. Raman scattering, generated with a 9 mW of 633 nm radiation, was collected for 60 s per window of spectrum using an 1800 grooves/mm diffraction grating. The LabSpec6 software recorded the Raman spectrum for each sample in the spectral range 400 cm^{-1} to 1800 cm^{-1} . The protocol for the extraction of the EVs from the cell culture supernatant using the TEIR reagent and conducting the SERS on the resuspended EVs in PBS is shown in Fig. 4.4.

^{‡‡} EVs extracted by Dr. Sreedevi P, Autophagy lab, JNCASR

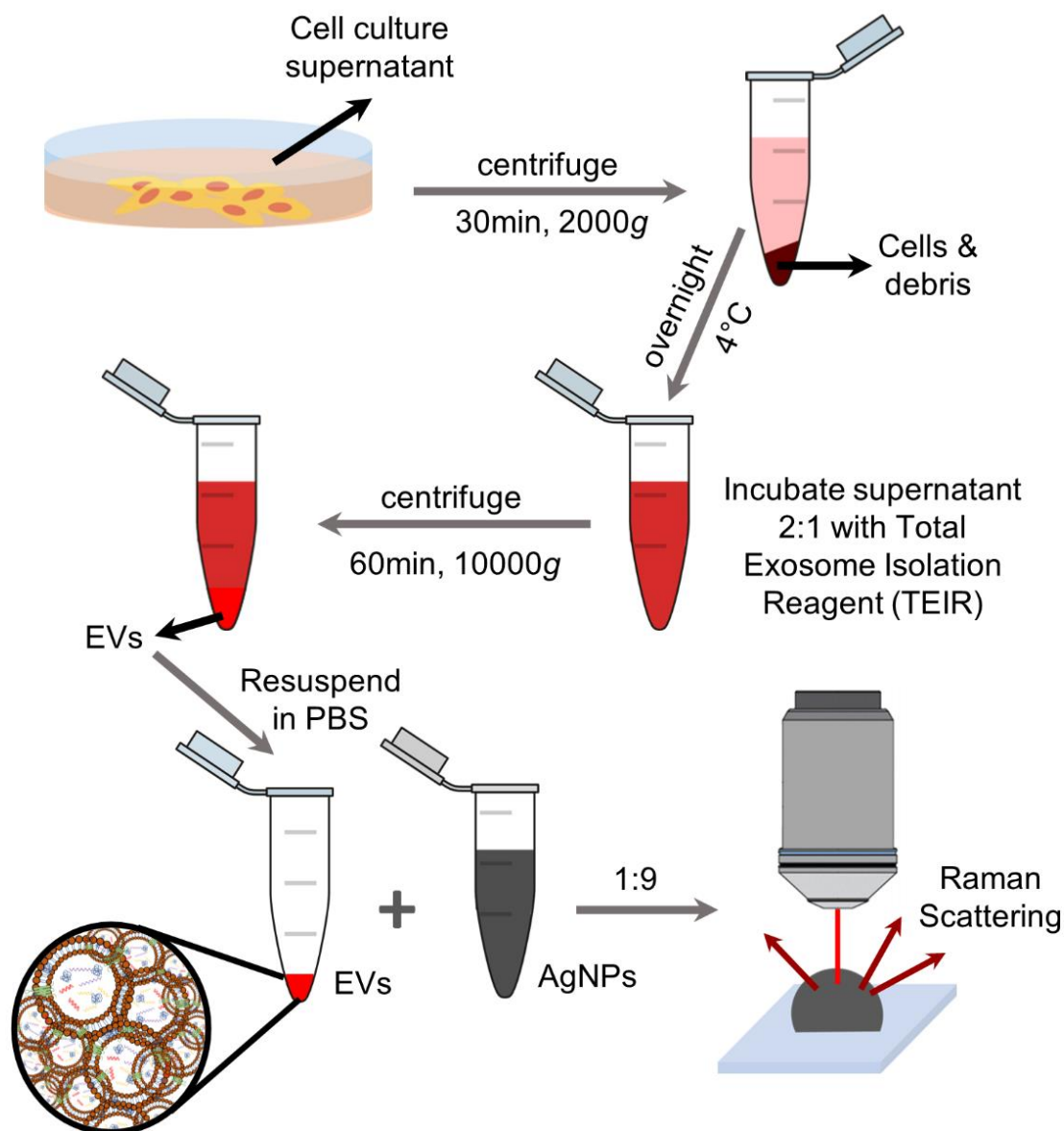


Fig. 4.4. Schematic representation of the experimental protocol.

c) Raman spectral analysis:

On the LabSpec6 software, the cosmic radiation spikes were removed. A user-defined baseline was used to correct the spectra on the MATLAB software (Fig. 4.5). After baseline subtraction, the spectra were normalized by setting the total area under the curve to one. The Raman spectral peaks were fit with a Lorentz curve, and the parameters thus obtained were used to perform the necessary calculations. PCA of the spectra was performed using a built-in MATLAB function, and then the first and second principal components were used for comparison.

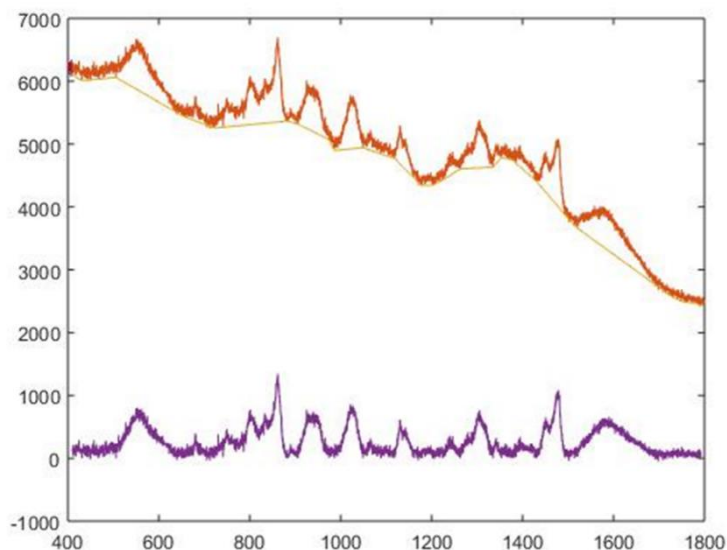


Fig. 4.5. User-defined baseline on MATLAB to correct all the spectra before they are area-normalized to equal a total area under the curve to one.

4.3. Results and Discussion

4.3.1. Western Blot and TEM characterisation

EVs are generally isolated from a serum-free medium.³² However, it is observed that cellular starvation, to various degrees, induces autophagy, which would result in a more vesicular release from the cells through endosomal autophagy and secretory autophagy.^{33, 34} The cells of HeLa, *Atg5*^{-/-} and HEK293 cell lines were subjected to autophagy induction by nitrogen starvation (EBSS) treatment. The EVs extracted from these cells under these starvation conditions using TEIR were analysed using western blot to detect the presence/absence of ATG5, CD63, and GM130 (Fig. 4.6). The western blot images clearly show that the *Atg5* protein is not expressed in the *Atg5*^{-/-} cell line and the EVs (as they needn't carry it). The anti-rabbit GM130 marker acts as a cytosolic (an intracellular marker) and is supposed to be only expressed in cells and absent in EVs; it is replicated accurately in the western blot. The CD63 protein is expressed in the cell line as a thick dark band and appears in the EVs on long/ over-exposure, which is also accurate as the cells contain large quantities of this protein compared to EVs.

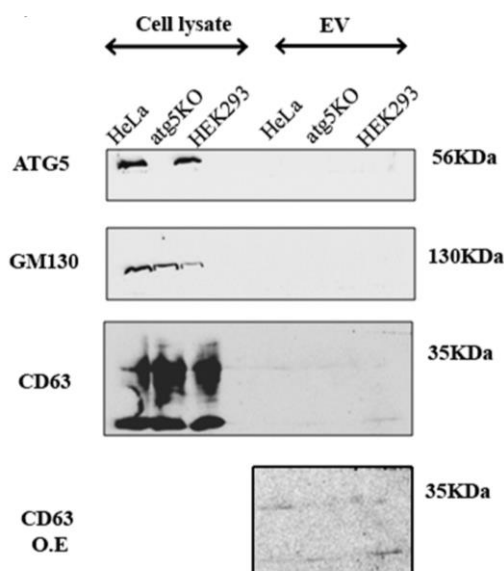


Fig. 4.6. Western blot analysis of the EVs isolated from HeLa, $Atg5^{-/-}$ and HEK293 cell lines.

TEM images of the EVs from HeLa and $Atg5^{-/-}$ cells and HeLa EV with AgNPs were collected using the JEOL 200 keV system in normal and cryo modes (Fig. 4.7).

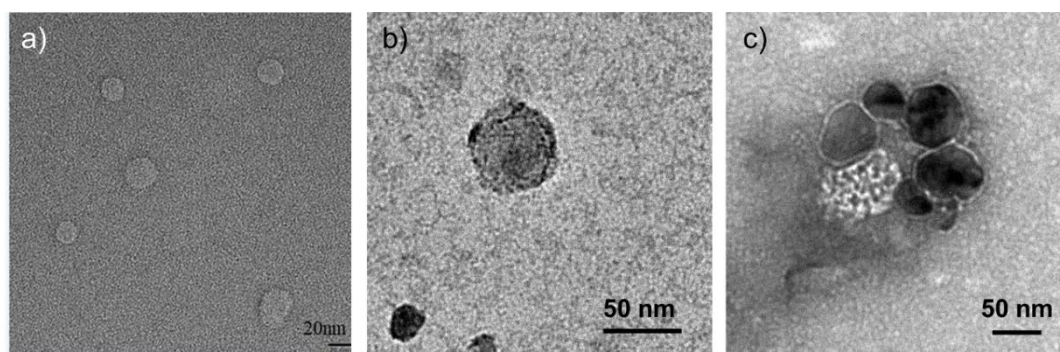


Fig. 4.7. a) Cryo-TEM image of the EV from the HeLa, b) Normal TEM image at 200 keV of the EV from the $Atg5^{-/-}$ cell line, c) Normal TEM image of a HeLa EV surrounded by AgNPs.

4.3.2. Surface-Enhanced Raman Spectroscopy

To collect the SERS spectra of the EVs, nine parts of concentrated AgNPs, and one part of EVs resuspended in PBS were mixed and incubated on ice. From the TEM images (Fig. 4.7c), we can get an idea that a few nanoparticles in the order of 2-5 can surround one EV. The ratio selection was made empirically, where we provide maximum opportunity for the EVs to be surrounded by the AgNPs, thus creating larger hotspot possibilities and improving the SERS enhancement. In general, it is known that the citrate capped AgNPs are electrochemically stabilized, and an increase in the concentration of a positive ion (Na^+ in the case of PBS) results in a screening

of surface charge, thus causing the AgNPs to stick and form agglomerates. From the DLS measurements, it can be seen very clearly that the AgNPs, have various size ranges, with a maximum number of them having a hydrodynamic radius of around 150 nm. In comparison, the AgNPs and PBS mixture incubated on ice for 8min shows a larger size distribution, with a maximum around 800 nm indicating aggregates of AgNPs (Fig. 4.8). The aggregation of the AgNPs as the buffer concentration increases in the nanoparticle EVs mixture is primarily due to the sorption of phosphate ions onto the surface of the nanoparticles.^{35, 36} It is plausible that there could be a regulated interaction of EVs with the SERS hotspots by incubating the mixture on ice (Fig. 4.9).

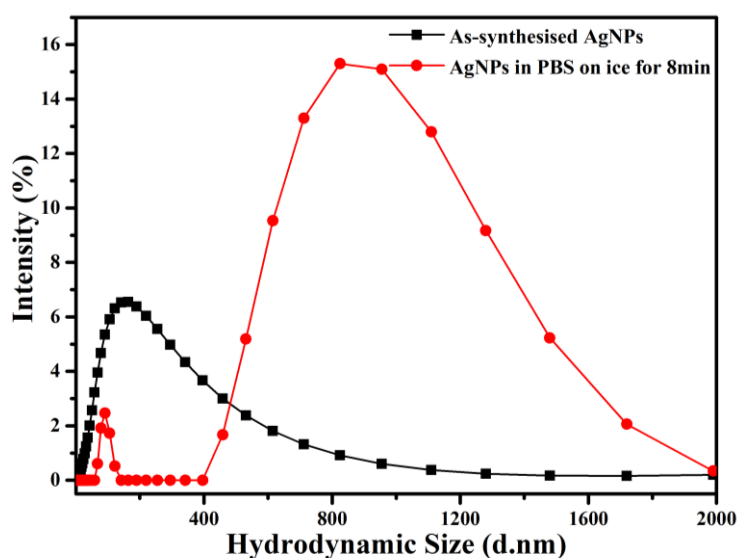


Fig. 4.8. DLS measurements of the AgNPs and a mixture of AgNPs in PBS on ice for 8min.

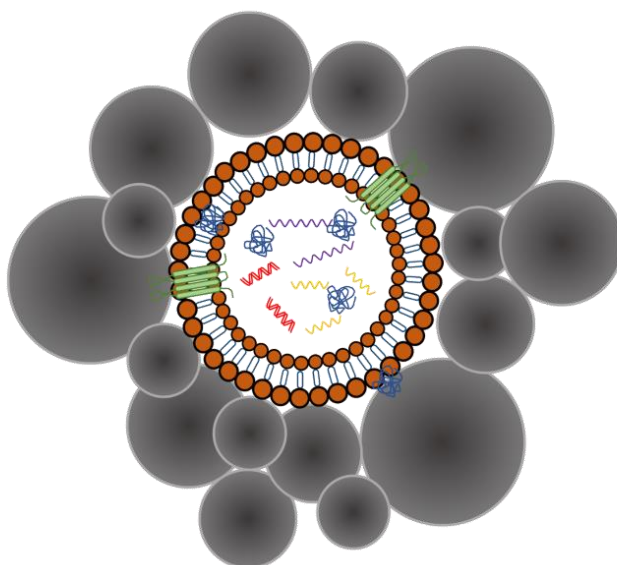


Fig. 4.9. Schematic representation of the EV surrounded by AgNPs during SERS.

Laser tweezers Raman spectroscopy (LTRS) studies in individual exosomes from eight-cell lines revealed that the chemical composition of the exosomes could be studied, but the cell line averaged spectra would vary considerably.³⁷ EVs were extracted from three different sets of culture supernatants for each cell line in these experiments. In the sample under the micron laser spot size, there would be thousands of EVs excited by the incident laser light, and hence every obtained Raman spectra is an average spectrum obtained as a result of a large number of EVs. The SERS spectra were obtained for TEIR as control, HeLa EVs, HEK293 EVs, and Atg5^{-/-} EVs in the wavenumber range of 400 to 1800 cm⁻¹. After subtracting the background using a user-defined baseline, the spectra were area-normalized to one. The reproducibility of these SERS spectra for the EVs provided consistent results (Fig. 4.10). The consistent results obtained for about 15 SERS spectra over 3 experiments were collected from EVs of different cell lines, and their averages were obtained. The characteristic peaks of the SERS spectra that help distinguish EVs from the different cell lines are represented in Fig. 4.11.

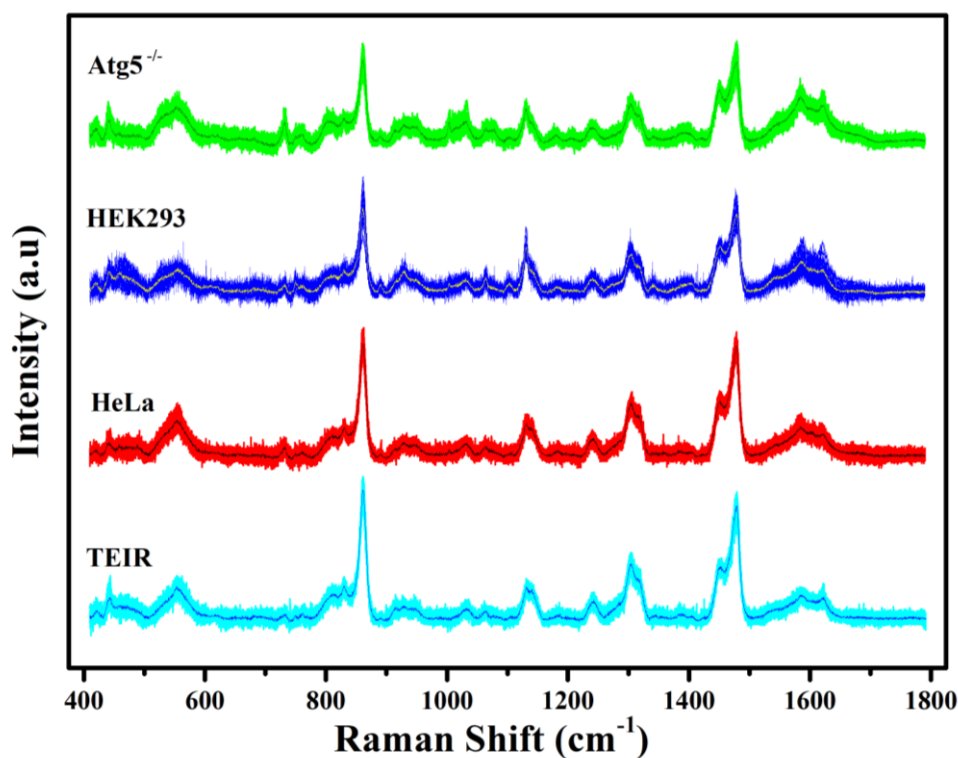


Fig. 4.10. Consistent result of the 15 SERS spectra of the EVs from different cell lines. The contrasting line in the middle of the spectra distribution is the sum average of the 15 spectra.

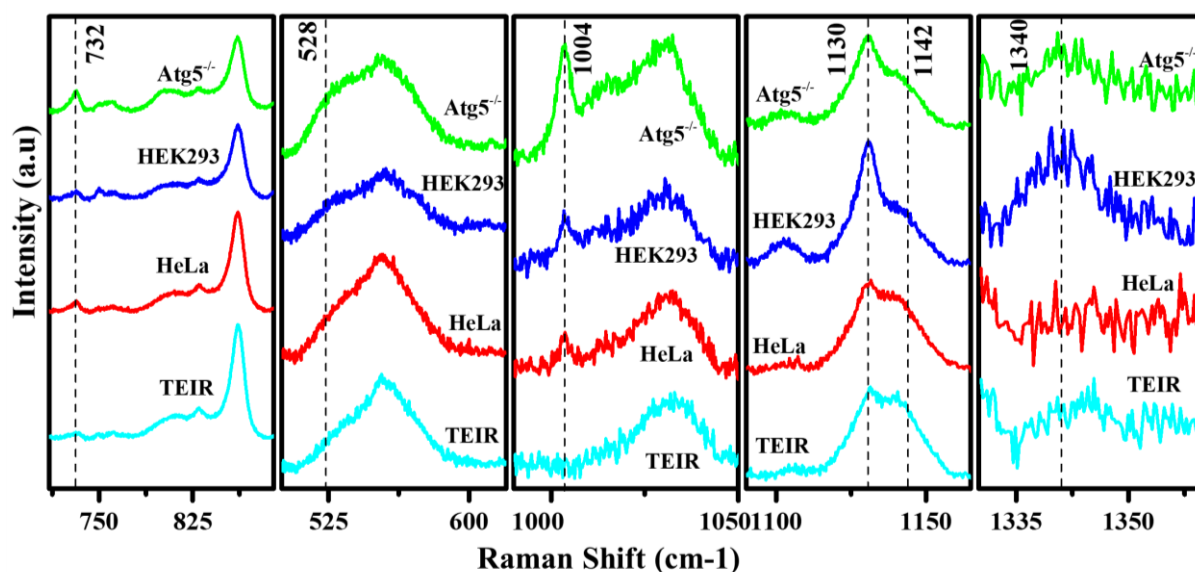


Fig. 4.11. Representative SERS spectra of the EVs from the HeLa, HEK293, and $Atg5^{-/-}$ cell lines and a reference spectrum of the TEIR. The characteristic peaks of the average SERS spectra collected for about 15 samples show differences amongst EVs extracted from different cell lines (Cyan – TEIR, Red – HeLa, Blue – HEK293, Green – $Atg5^{-/-}$).

In general, SERS produces a stronger enhancement of Raman signals from the parts of the EV closer to the nanoparticles' surface than its components, which are farther away from them. Also, at any instance of collecting the SERS, the resulting spectrum was an average from all the EVs illuminated in the laser volume. Since major peaks in the SERS spectra arose from the TEIR itself, difference spectra of EVs from HeLa, HEK293, and $Atg5^{-/-}$ cell lines to TEIR were plotted (Fig. 4.12a-c). The difference plots between SERS of EVs from HeLa–HEK293, HeLa– $Atg5^{-/-}$ and $Atg5^{-/-}$ –HEK293 were plotted (Fig. 4.12d-f). The difference plots, as observed in Fig. 4.12 d-f give us an idea of the changes in the chemical constituents of the cell lines since the SERS spectra manifest the differences in the vibrational modes of the system due to the compositional variation of EVs. These changes are due to the differences in lipids, proteins, nucleic acids on the EVs surfaces.

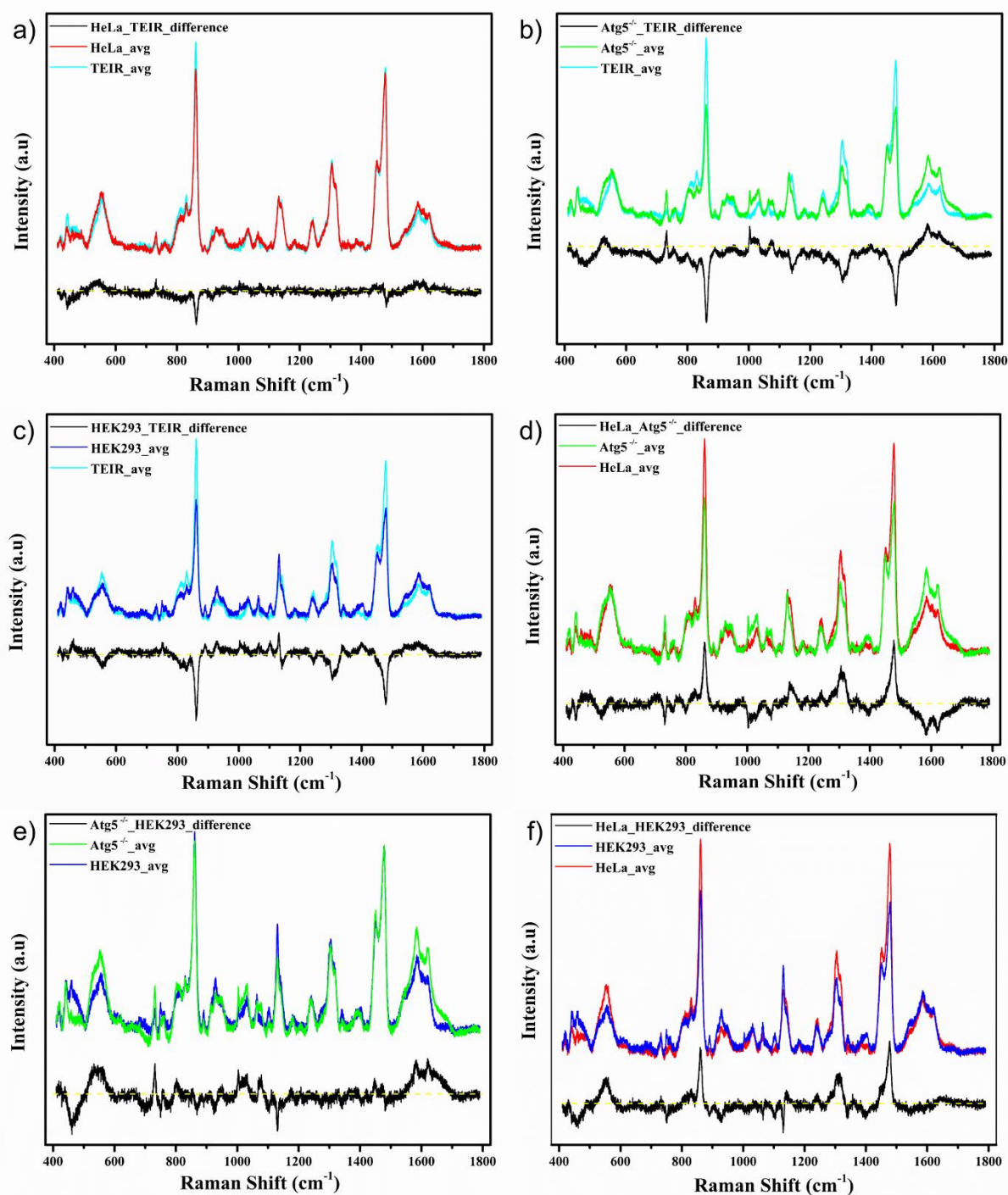


Fig. 4.12. A comprehensive profile of difference plots was obtained by subtracting the averaged SERS spectra. Difference plots were obtained for a) EVs from HeLa and TEIR, b) EVs from Atg5^{-/-} and TEIR, c) EVs from HEK293 and TEIR. d) EVs from HeLa and Atg5^{-/-} cell lines e) EVs from Atg5^{-/-} and HEK293 cell lines f) EVs from Atg5^{-/-} and HEK293 cell lines.

4.3.2.1. Biochemical analysis of the SERS spectra

The 732 cm⁻¹ Raman peak, corresponding to phosphatidylserine,³⁸ was prominent in the EVs from HeLa and the Atg5^{-/-} cell lines but not found in HEK293 (Fig. 4.11). It has been often reported that phosphatidylserine is present on the surface

of EVs for cancerous cells and in the inner leaflets of membrane bilayer in healthy cells.³⁹ Enterovirus exit cells by exploiting phosphatidylserine positive vesicles originating from autophagosomes.^{40, 41} Exposure of phosphatidylserine on the exterior acts as an apoptotic signal.⁴² As the surface phosphatidylserine is responsible for the internalization of microvesicles derived from hypoxia-induced human bone marrow mesenchymal stem cells,⁴³ our results of elevated phosphatidylserine levels in the EVs isolated from the cancer cells of HeLa and its autophagy knockout, but lower in the HEK293 EVs are comparable. Hence, this peak can act as a SERS marker to differentiate healthy cells from cancer cells.

A shoulder peak around 528 cm^{-1} , corresponding to cholesterol and cholesteryl esters,^{38, 44} was found for the EVs from all three cell lines (Fig. 4.11). Many lines of evidence suggest marked differences in the enrichment of cholesterol, sphingomyelins, glycosphingolipids, and phosphatidylserine from cells to exosomes.⁴⁵ *In vitro* studies reveal that cholesterol enrichment helps the vascular smooth muscles and monocytes enhance the release of microvesicles.^{46, 47} Raman studies show that exosomes of non-cancerous cell lines have more cholesterol content when compared to those of cancerous cell lines.³⁷ Our results demonstrate considerable enrichment of cholesterol/cholesteryl esters in the *Atg5*^{-/-} EVs, which could plausibly mean that they favour vesicular fusion with the plasma membrane by controlling the content of specific proteins at sites of exocytosis by lipid regulation.⁴⁸ Although the link between cholesterol homeostasis and EV release is not well understood,⁴⁹ it has been reported that the cholesteryl esters (CEs), specifically CE (18:1), can discriminate between prostate cancer, non-tumour, and benign prostatic hyperplasia.^{50, 51} The abundance of cholesteryl esters is higher in metastatic prostate cancer PC-3 EVs than normal RWPE1 EVs and a drop in NB26 EVs. Moreover, cholesterol and cholesteryl esters share many common Raman peaks.⁵²

The 1004 cm^{-1} peak corresponds to phenylalanine (Fig. 4.11). It is observed that phenylalanine, which is the precursor of tyrosine, is enriched in exosome-eluted fractions.⁵³ It is reported that tyrosine dependent inactivation of mTORC1 and

starvation induces autophagy, which would probably help convert phenylalanine to tyrosine under autophagic conditions.⁵⁴ In an autophagy-deficient cell line under starvation conditions, the flux of conversion of phenylalanine to tyrosine might not be pronounced. This could be the possible reason why the phenylalanine peak at 1004 cm^{-1} appears comparatively higher in the EVs from *Atg5*^{-/-} compared to HeLa cell EVs.

A doublet of peaks exists in the TEIR at 1130 cm^{-1} and 1142 cm^{-1} (Fig. 4.11). The 1130 cm^{-1} peak also corresponds to fatty acids like tripalmitin, tristearin, triarachidin, tribehenin.⁴⁴ But the 1142 cm^{-1} peak only originates from TEIR. The change in the intensity ratio of these two peaks suggests the presence of fatty acids in EVs. Mounting lines of evidence indicate that fatty acids are abundant in EVs from non-tumorigenic cells, which is observed and comparable even in our SERS studies.⁵⁵

As nucleic acids and proteins are well-known components of EVs, our experiment confirms the same. A small peak at 1340 cm^{-1} corresponding to nucleic acids is seen in HEK293 and *Atg5*^{-/-} EVs (Fig. 4.11).⁵⁶ It is interesting to note that the nucleic acid band at 1340 cm^{-1} is not present in the EVs of HeLa but visible in the *Atg5*^{-/-}. It is reported that the removal of damaged self DNA by Dnase2a requires autophagy-mediated delivery of the DNA to lysosomes.⁵⁷ In conjunction with a recent observation that prevention of autophagy-lysosome fusion increases exosome secretion,⁵⁸ these notions imply that exosome secretion and autophagy may act in a complementary manner to remove pro-inflammatory DNA from cells.⁵⁹ It is probably because the nucleic acid peak at 1340 cm^{-1} is seen in autophagy-deficient EVs compared to HeLa, where they possibly get degraded within the lysosomes.

A careful analysis of various spectral aspects finds differences amongst the spectra. Overall, the analysis of these various SERS peaks amongst the EVs extracted from different cell lines provides us with a picture of the difference in the chemical composition amongst them. The Raman peaks and their assignments and behaviours amongst the cell lines are summarized in the following Table. 4.1.

Table. 4.1. Differences in the SERS spectra as observed amongst the EVs from the various cell lines and their chemical origins.

	TEIR	HeLa	HEK293	Atg5 ^{-/-}	Raman assignment	Ref
I ₅₂₈ /I ₅₅₅	--	2.9	3.1	3.4	528 cm ⁻¹ corresponds to cholesterol /s-s disulphide	38, 44
732		✓	✓ (small)	✓	phosphatidylserine	38
1004	--	1	1.2	3	phenylalanine	60, 61
I ₁₁₃₀ /I ₁₁₄₂	1.1	1.1	2.7	2.4	1130 cm ⁻¹ from tripalmitin/tristearin/triarachidin/tribehenin	44
I ₁₃₄₀			✓	✓	Nucleic acid modes content differences	56

4.3.3. Principal Component Analysis

As we have seen so far, many of the peaks were common with the TEIR reagent; hence it was not straightforward to mark out the differences amongst the spectra of the EVs. We had to collect the difference spectra, and plot out the regions that were not arising from the nanoparticles, and hunt for subtle differences amongst the spectra. This effort can be practical in understanding the chemical origin differences and scientifically interpreting the differences between the EVs extracted from different cell lines. As the task is to differentiate the EVs, a simpler dimensionality reducing tool like PCA can help put the whole scenario in one picture and easily differentiate them. As shown in Fig. 4.13, the principal component 1 (PC1) vs. the principal component 2 (PC2) was made among the EVs from different cell lines and TEIR itself. As shown in Fig. 4.13, the cyan circles correspond to the TEIR, red triangles to HeLa EVs, blue squares to HEK293 EVs, and the green stars to the Atg5^{-/-} HeLa EVs. The clusters of the various PC components have been represented by ellipses, with the following corresponding equation,

$$\frac{((x - h) \cos \theta + (y - k) \sin \theta)^2}{a^2} + \frac{((x - h) \sin \theta + (y - k) \cos \theta)^2}{b^2} = 1 \quad (4.1)$$

Where $2a$ represents the length of the major axis and $2b$ the length of the minor axis (when $a > b$); (h, k) represents the centre of the ellipse, and θ (in rad) is the angle of the major axis measured counter-clockwise from the x-axis. The ellipses corresponding to the various clusters can be represented by this equation with the set of parameters (a, b, h, k, θ) valued at $(0.1443, 0.1154, 0.0114, 0.0377, 0.1479)$ for TEIR; $(0.1396, 0.0346, 0.0106, 0.0674, 0.0428)$ for HeLa; $(0.1081, -0.0148, 0.0140, 0.1989, -0.0216)$ for HEK293; and $(0.1097, 0.1780, 0.0217, 0.0611, 0.0466)$ for $Atg5^{-/-}$. The distinct clusters formed in the plot show us that PCA can differentiate the variations in the SERS spectra of the various EVs. Of the two blind experiments performed, one was predicted as EVs from the HeLa cell line and was accurate to the actual sample provided. The other was predicted as either TEIR or EVs from the HeLa cell line, as the point fell in the intersection region of the two, while the sample provided was EVs from the HeLa cell line. Based on the blinded study, we could predict the cell line from where the EVs were extracted with a good probability (Fig. 4.13).

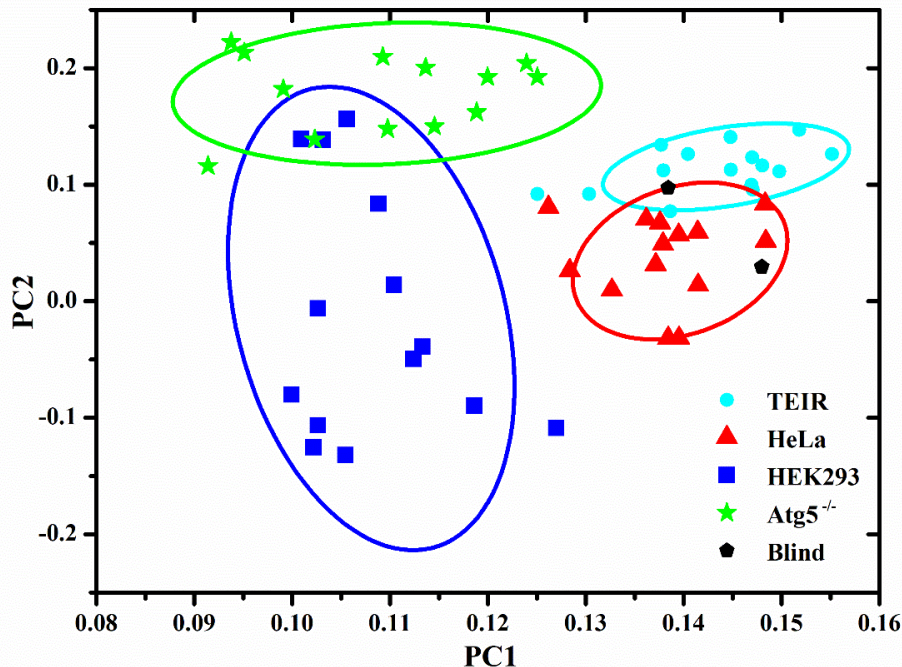


Fig. 4.13. Scoreplot of PCA for SERS spectra of TEIR; EVs from HeLa, HEK293 and $Atg5^{-/-}$ cell lines and a blinded study. The data points are marked as cyan dots for TEIR, red triangles for HeLa EVs, blue squares for HEK293 EVs, green stars for $Atg5^{-/-}$ EVs, and black pentagons for EVs from blinded studies.

So far, we have established using SERS combined with PCA to differentiate EVs extracted from different cell lines. We chose another set of cell lines to test the strategy and extracted the EVs from them. For a good comparison with previous results, we chose the following cell lines, SHSY5Y (human-derived neuroblastoma cell line - cancerous), HaCaT (cultured human keratinocyte cell line - noncancerous) and RAW264.7 macrophages (mouse origin cancerous cell line). The EVs from these cell lines were also extracted under autophagic conditions using TEIR. The SERS spectra of these EVs were collected, baseline-corrected and area-normalized for further processing. A PCA plot depicting the principal components with 75% confidence circles was constructed to compare the EVs from these cell lines.

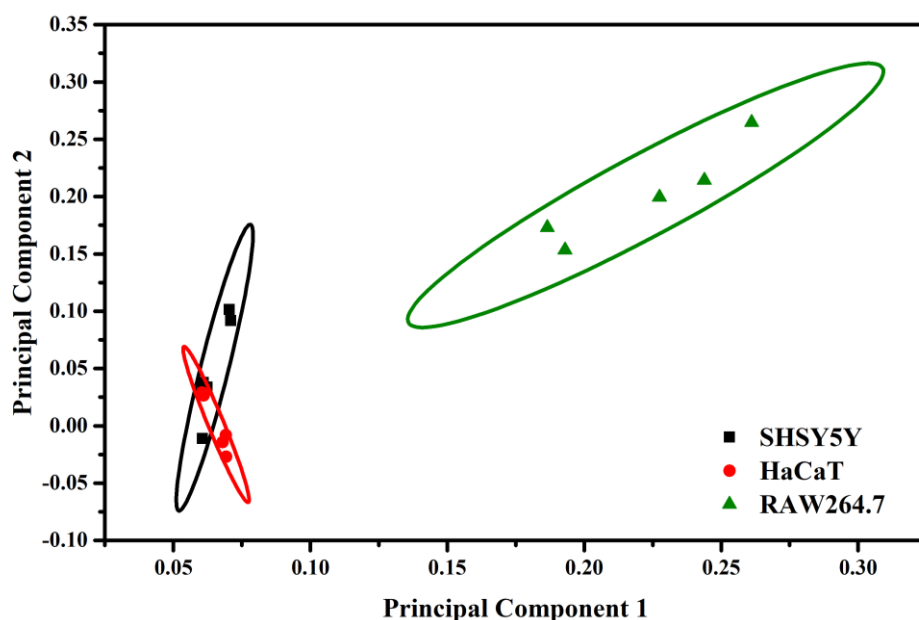


Fig. 4.14. Scoreplot of PCA for SERS of EVs extracted from SHSY5Y, HaCaT, and RAW264.7 macrophage cell lines.

As seen in the scoreplot in Fig. 4.14, the SERS spectra of the different EVs extracted from different cell lines have formed unique clusters. The two human origin cell lines form two overlapping clusters, while the mouse origin cell line is a very distinct cluster. If the PC1 and PC2 components for the human cell lines EVs are separately plotted, the differences amongst them can also be clearly distinguished. Biochemical analysis and interpretation would be required to understand the overlaps and differences further. From these studies, we have established that the SERS combined with PCA can be used to differentiate EVs with a good amount of confidence.

4.3.4. SERS for EVs extracted by ultracentrifugation method

Thus far, we have looked at the EVs extracted from various cell lines using the isolation reagent TEIR. The major drawback was that most of the SERS spectrum was from the reagent itself, and difference plots had to be made to extract the subtle bands corresponding to the EVs. Though easy to conduct experimentally, this method reduces the percentage confidence of using the technique for actual detection. It is still reliable with a large number of data points. To test out the SERS based PCA detection, we conducted the same set of experiments using exosomes extracted from some cell lines *via* the classical ultracentrifugation method.

By this extraction method, varying centrifugation speeds at each step lead to the precipitation of different components, including cells, dead cells, cell debris, and EVs of different sizes. Our present study used two cell lines, HEK293T (represented as HEK) and HEK293T cells transfected with full-length SHH (represented as Shh). Shh is a sonic hedgehog signalling protein for regulating embryonic morphogenesis in animals. Exosomes were extracted using the ultracentrifugation method at $150,000g$ and $450,000g$ and are represented as HEK_150, HEK_450, Shh_150 and Shh_450. The SERS spectra were recorded thrice to maintain consistency (Fig. 4.15a). The processed SERS spectra were used to obtain the PCA plot (Fig. 4.15b).^{§§}

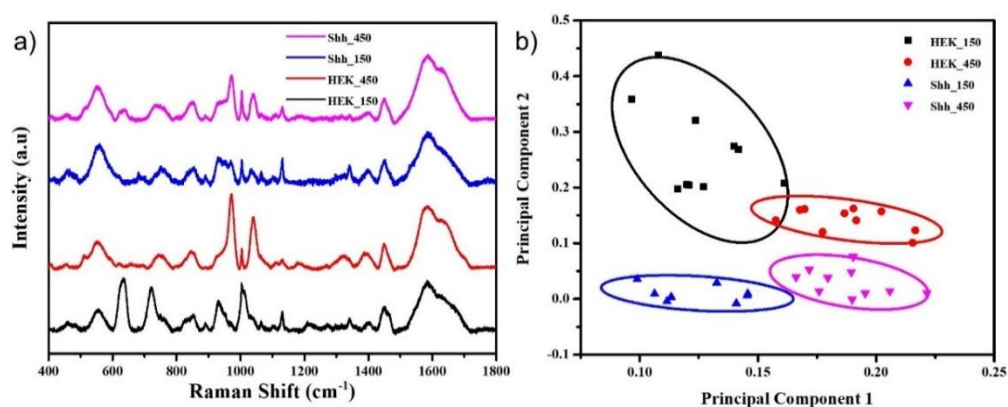


Fig. 4.15. a) Average SERS spectra of EVs from HEK293T and SHH transfected HEK293T cells, extracted at $150,000g$ and $450,000g$, b) PCA scoreplot for the collected SERS spectra.

^{§§} Exosome samples provided by Dr. Neha Vyas's group, St. John's hospital, Bengaluru

The collected SERS spectra visibly showed distinct differences, consistent with the scoreplot for the PCA of the SERS spectra. Hence, confirming that the strategy works for EVs extracted by any technique.

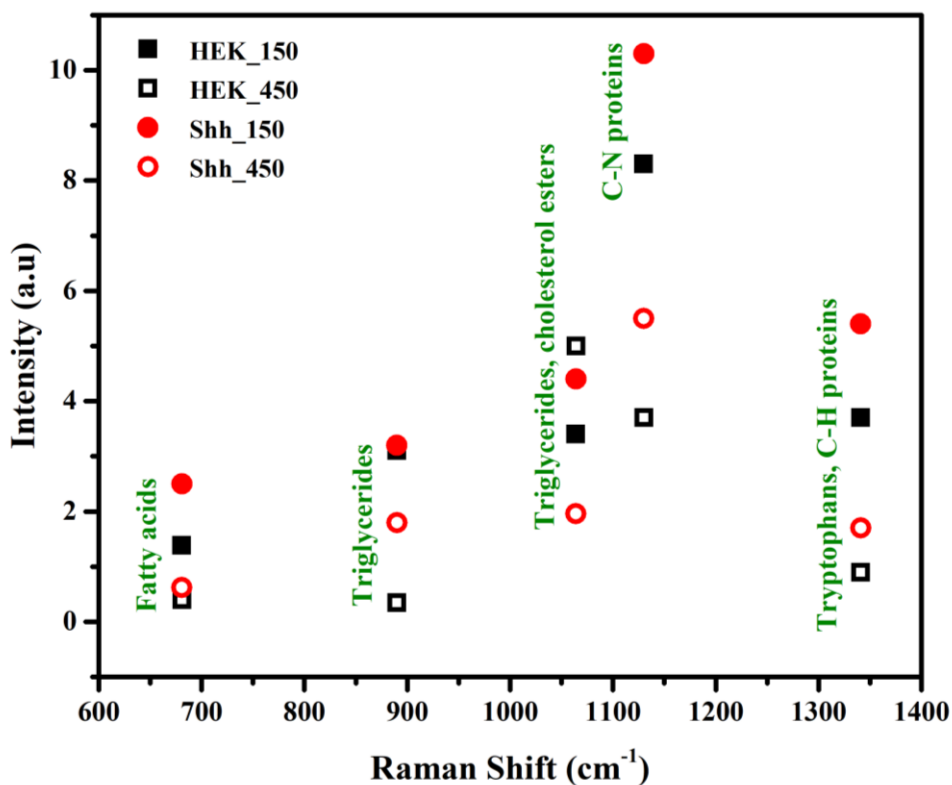


Fig. 4.16. Peak intensities in the SERS spectra of the various exosomes extracted from HEK293T and SHH transfected HEK293T cell lines.

From Fig. 4.16, it can be seen that the peaks around 679 cm^{-1} , 889 cm^{-1} and 1064 cm^{-1} are higher in the Shh_150 over the HEK_150, indicating a higher percentage of triglycerides and cholesterol esters in the exosomes extracted from the SHH transfected HEK293T over the ones extracted from HEK293T cell line. The 679 cm^{-1} and 889 cm^{-1} peaks are higher for the Shh_450 over the HEK_450, while *vice versa* in the case of 1064 cm^{-1} , indicating differences in the chemical compositions of the four exosomes. Further lipidomics, proteomics and genomics study can provide compositional insights into these exosomes, hence understanding the Raman spectra better.

Advances in understanding the composition of the EVs provide more references in interpreting the Raman spectra, which can further be used as a simple sorting and

interpreting technique. As of now, Raman spectroscopy has been established to differentiate EVs extracted from various cell lines.

The following Table. 4.2 provides a comparison of the strategy in this work compared to the many other techniques for detecting EVs. From this table, the PCA analysis of the SERS is seen to be a quick technique used to differentiate EVs, which needs small sample volumes and can provide a few insights into the molecular compositions.

Table. 4.2. Comparative summary of the properties of various approaches for EV analysis and detection.

Parameters	Measurement type	Size Detection & Size Range	Sample Processing/ Ease of experiment	Measurement time	EVs vol. req.	Molecular composition
<i>Biochemical Analysis</i>						
Western Blot	Bulk	NA	+	*	large	yes
ELISA	Bulk	NA	+	*	large	yes
<i>Surface Plasmon Resonance Based Technique</i>						
SERS (Label-free)	Bulk	NA	+++	***	small	yes
SERS (specific)	Bulk	NA	+++	***	small	yes
LSPRi	Bulk/ Individual	NA	++	**	small	yes
<i>Direct Imaging Technique</i>						
TEM	Individual	> 5 nm	++	**	small	no
AFM	Individual	> 5 nm	++	**	small	no
<i>Indirect Optical Imaging Technique</i>						
DLS	Bulk	5-2000 nm	+++	***	large	no
NTA	Individual	50-1000 nm	+++	***	small	no
<i>Indirect non-Optical Imaging Technique</i>						
TRPS	Individual	> 30nm	++	**	small	yes

+ - difficult to +++ - easy; *- time consuming to ***- quick results, NA-not applicable (these techniques cannot be used to obtain size parameters)

4.4. Conclusions

We have conducted the SERS of EVs extracted from the three cell lines, including HeLa, HEK293, Atg5^{-/-}, and used this to differentiate EVs extracted from them. We can see a good correlation between the observed SERS spectra and the occurrence of nucleotides and metabolites in EVs from different cell lines. The SERS spectra of the EVs from these cell lines were easily distinguished using PCA, where each one of them formed clear clusters in the score plot. Our study helps identify the components of EVs by the differences in spectra of the EVs secreted by the cell lines that are very closely associated with each other, as seen in the case of HeLa and Atg5^{-/-}. We could also study exosomes extracted by the ultracentrifugation method using SERS and PCA analysis, and obtain different clusters, indicating the ability of the technique to differentiate EVs extracted by different techniques. As exosome specific antibody-decorated SERS nanoprobe have been employed to differentiate EVs from tumour cells and healthy cells,⁶² EVs are used as diagnostic markers in cancer.⁶³

Proteomic studies in the EVs isolated from human breast cancer cells helps in the stratification of different cancer subtypes amidst their wide heterogeneity. Cell-derived EV subtyping in glioblastoma cells is also reported.⁶⁴ Our method employs the simple precipitation procedures of EV isolation with a heterogeneous population and can distinguish EVs from different cell lines (HeLa and HEK293) and between two closely related cell lines of HeLa. This would, in turn, aid in establishing a repository to differentiate EVs from various cell lineages. In most of the reported literature, ultracentrifugation procedures are performed, whereas our assay employs a simple precipitation-based method that can be used for quick analysis. Our study demonstrates a more simplified usage of SERS in analysing EVs from different cell lines. Despite the high spectral variability reported in the EVs of the same cell type using LTRS, our study employing SERS seems more robust in differentiating EVs from different cell lines. This could be a good screening tool for disease diagnostics in the future, coupled with other techniques. Our study warrants the need to develop more robust techniques to understand the biogenesis and composition of EVs.

Further, this study employs minimal sample volume with even lesser protein concentrations, which are not detectable in the western blots. Such studies would help develop potent diagnostic markers of biological fluids where the sample size is limiting. Advancements in technology and methodology may lead to a renewed understanding and definition of various subtypes of EVs, including exosomes. Novel discoveries continue to add dimensions of complexity into evolving models of exosome–autophagy interactions and vesicular trafficking. SERS is one such technique that, when combined with other techniques such as AFM, can be used as a powerful tool to understand the composition of the EVs better.⁶⁵

4.5. References

1. Jeppesen, D.K. et al. Reassessment of Exosome Composition. *Cell* 177, 428 (2019).
2. Kalra, H., Drummen, G.P.C. & Mathivanan, S. Focus on Extracellular Vesicles: Introducing the Next Small Big Thing. *International journal of molecular sciences* 17, 170 (2016).
3. Théry, C. Exosomes: secreted vesicles and intercellular communications. *F1000 biology reports* 3, 15 (2011).
4. Théry, C. et al. Minimal information for studies of extracellular vesicles 2018 (MISEV2018): a position statement of the International Society for Extracellular Vesicles and update of the MISEV2014 guidelines. *Journal of extracellular vesicles* 7, 1535750 (2018).
5. Lobb, R.J. et al. Optimized exosome isolation protocol for cell culture supernatant and human plasma. *Journal of Extracellular Vesicles* 4, 27031 (2015).
6. Théry, C., Ostrowski, M. & Segura, E. Membrane vesicles as conveyors of immune responses. *Nature Reviews Immunology* 9, 581 (2009).
7. Costa-Silva, B. et al. Pancreatic cancer exosomes initiate pre-metastatic niche formation in the liver. *Nature Cell Biology* 17, 816 (2015).
8. Zappulli, V., Friis, K.P., Fitzpatrick, Z., Maguire, C.A. & Breakefield, X.O. Extracellular vesicles and intercellular communication within the nervous system. *The Journal of Clinical Investigation* 126, 1198 (2016).
9. Chuo, S.T.-Y., Chien, J.C.-Y. & Lai, C.P.-K. Imaging extracellular vesicles: current and emerging methods. *Journal of Biomedical Science* 25, 91 (2018).
10. Zhang, Y., Liu, Y., Liu, H. & Tang, W.H. Exosomes: biogenesis, biologic function and clinical potential. *Cell & Bioscience* 9, 19 (2019).
11. Khandia, R. et al. A Comprehensive Review of Autophagy and Its Various Roles in Infectious, Non-Infectious, and Lifestyle Diseases: Current Knowledge and Prospects for Disease Prevention, Novel Drug Design, and Therapy. *Cells* 8, 674 (2019).
12. Bhutia, S.K., Mukhopadhyay, S., Sinha, N., Das, D.N., Panda, P.K., Patra, S.K., Maiti, T.K., Mandal, M., Dent, P., Wang, X.Y., Das, S.K., Sarkar, D., Fisher, P.B. Autophagy: cancer's friend or foe? *Advances in Cancer Research* 118, 61 (2013).

13. Yun, C.W. & Lee, S.H. The Roles of Autophagy in Cancer. *International Journal of Molecular Sciences* 19 (2018).
14. Garcia, N.A., Ontoria-Oviedo, I., González-King, H., Diez-Juan, A., Sepúlveda, P. Glucose Starvation in Cardiomyocytes Enhances Exosome Secretion and Promotes Angiogenesis in Endothelial Cells. *PLOS ONE* 10, e0138849 (2015).
15. Guo, W. et al. Exosomes: New players in cancer (Review). *Oncology reports* 38, 665 (2017).
16. Guo, H., Chitiprolu, M., Roncevic, L., Javalet, C., Hemming, F. J., Trung, M. T., Meng, L., Latreille, E., Tanese de Souza, C., McCulloch, D., Baldwin, R. M., Auer, R., Côté, J., Russell, R. C., Sadoul, R., Gibbings, D. Atg5 Disassociates the V(1)V(0)-ATPase to Promote Exosome Production and Tumor Metastasis Independent of Canonical Macroautophagy. *Developmental Cell* 43, 716 (2017).
17. Cizmar, P., Yuana, Y. in *Extracellular Vesicles-Methods in molecular biology* (ed. Kuo W., J.S.) (Humana Press, New York, 2017).
18. Rikkert, L. et al. Cancer-ID: Toward Identification of Cancer by Tumor-Derived Extracellular Vesicles in Blood. *Frontiers in Oncology* 10 (2020).
19. Park, J. et al. Exosome Classification by Pattern Analysis of Surface-Enhanced Raman Spectroscopy Data for Lung Cancer Diagnosis. *Analytical Chemistry* 89, 6695 (2017).
20. Lee, C., Carney, R., Lam, K. & Chan, J.W. SERS analysis of selectively captured exosomes using an integrin-specific peptide ligand. *Journal of Raman Spectroscopy* 48, 1771 (2017).
21. Tirinato, L. et al. SERS Analysis On Exosomes Using Super-Hydrophobic Surfaces. *Microelectronic Engineering* 97 (2012).
22. Im, H., Yang, K., Lee, H., Castro, C. M. in *Extracellular Vesicles. Methods in Molecular Biology* (ed. Kuo W., J.S.) (Humana Press, New York, 2017).
23. Coumans, F.A.W., Gool, E.L. & Nieuwland, R. Bulk immunoassays for analysis of extracellular vesicles. *Platelets* 28, 242 (2017).
24. Au - Skliar, M. & Au - Chernyshev, V.S. Imaging of Extracellular Vesicles by Atomic Force Microscopy. *JoVE*, e59254 (2019).
25. Palmieri, V. et al. Dynamic light scattering for the characterization and counting of extracellular vesicles: a powerful noninvasive tool. *Journal of Nanoparticle Research* 16 (2014).

26. Bachurski, D. et al. Extracellular vesicle measurements with nanoparticle tracking analysis - An accuracy and repeatability comparison between NanoSight NS300 and ZetaView. *Journal of extracellular vesicles* 8, 1596016 (2019).
27. Welsh, J.A., Holloway, J.A., Wilkinson, J.S. & Englyst, N.A. Extracellular Vesicle Flow Cytometry Analysis and Standardization. *Frontiers in Cell and Developmental Biology* 5 (2017).
28. Coumans, F.A.W. et al. Reproducible extracellular vesicle size and concentration determination with tunable resistive pulse sensing. *Journal of extracellular vesicles* 3, 25922 (2014).
29. Han, C. et al. Single-vesicle imaging and co-localization analysis for tetraspanin profiling of individual extracellular vesicles. *Journal of Extracellular Vesicles* 10, e12047 (2021).
30. Merdalimova, A.C., V.; Nozdriukhin, D.; Rudakovskaya, P.; Gorin, D.; Yashchenok, A. Identification and Analysis of Exosomes by Surface-Enhanced Raman Spectroscopy. *Applied Sciences* 9, 1135 (2019).
31. Shin, H., Jeong, H., Park, J., Hong, S. & Choi, Y. Correlation between Cancerous Exosomes and Protein Markers Based on Surface-Enhanced Raman Spectroscopy (SERS) and Principal Component Analysis (PCA). *ACS Sensors* 3, 2637 (2018).
32. Gudbergsson, J.M., Johnsen, K.B., Skov, M.N. & Duroux, M. Systematic review of factors influencing extracellular vesicle yield from cell cultures. *Cytotechnology* 68, 579 (2016).
33. Mejlvang, J. et al. Starvation induces rapid degradation of selective autophagy receptors by endosomal microautophagy. *Journal of Cell Biology* 217, 3640 (2018).
34. Ponpuak, M. et al. Secretory autophagy. *Current Opinion in Cell Biology* 35, 106 (2015).
35. Afshinnia, K. & Baalousha, M. Effect of phosphate buffer on aggregation kinetics of citrate-coated silver nanoparticles induced by monovalent and divalent electrolytes. *Science of The Total Environment* 581-582, 268 (2017).
36. White, P. & Hjortkjaer, J. Preparation and characterisation of a stable silver colloid for SER(R)S spectroscopy. *Journal of Raman Spectroscopy* 45, 32 (2014).

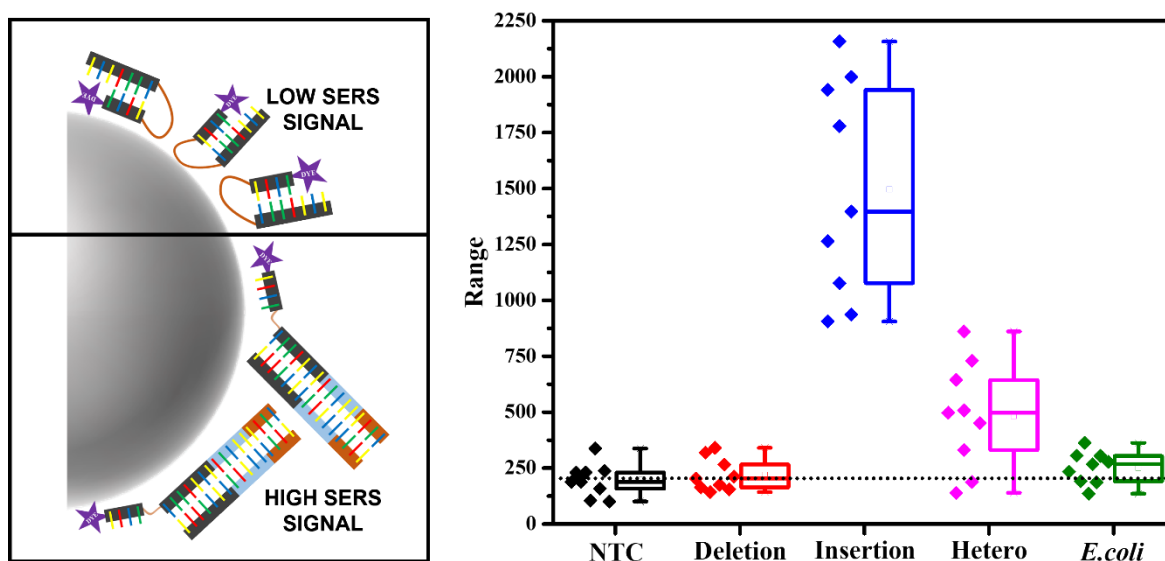
37. Smith, Z.J. et al. Single exosome study reveals subpopulations distributed among cell lines with variability related to membrane content. *Journal of Extracellular Vesicles* 4, 28533 (2015).
38. Krafft, C., Neudert, L., Simat, T. & Salzer, R. Near infrared Raman spectra of human brain lipids. *Spectrochimica acta. Part A, Molecular and biomolecular spectroscopy* 61, 1529 (2005).
39. N'Guessan, K.F., Patel, P.H. & Qi, X. SapC-DOPS – a Phosphatidylserine-targeted Nanovesicle for selective Cancer therapy. *Cell Communication and Signaling* 18, 6 (2020).
40. Chen, Y.-H. et al. Phosphatidylserine Vesicles Enable Efficient En Bloc Transmission of Enteroviruses. *Cell* 160, 619 (2015).
41. Mutsafi, Y., & Altan-Bonnet, N. Enterovirus Transmission by Secretory Autophagy. *Viruses* 10, 139 (2018).
42. Segawa, K. & Nagata, S. An Apoptotic ‘Eat Me’ Signal: Phosphatidylserine Exposure. *Trends in Cell Biology* 25, 639 (2015).
43. Wei, X. et al. Surface Phosphatidylserine Is Responsible for the Internalization on Microvesicles Derived from Hypoxia-Induced Human Bone Marrow Mesenchymal Stem Cells into Human Endothelial Cells. *PLOS ONE* 11, e0147360 (2016).
44. Czamara, K. et al. Raman spectroscopy of lipids: a review. *Journal of Raman Spectroscopy* 46, 4 (2015).
45. Skotland, T., Sandvig, K. & Llorente, A. Lipids in exosomes: Current knowledge and the way forward. *Progress in Lipid Research* 66, 30 (2017).
46. Liu, M.-L., Reilly Michael, P., Casasanto, P., McKenzie Steven, E. & Williams Kevin, J. Cholesterol Enrichment of Human Monocyte/Macrophages Induces Surface Exposure of Phosphatidylserine and the Release of Biologically-Active Tissue Factor-Positive Microvesicles. *Arteriosclerosis, Thrombosis, and Vascular Biology* 27, 430 (2007).
47. Llorente-Cortés, V., Otero-Viñas, M., Camino-López, S., Llampayas, O. & Badimon, L. Aggregated Low-Density Lipoprotein Uptake Induces Membrane Tissue Factor Procoagulant Activity and Microparticle Release in Human Vascular Smooth Muscle Cells. *Circulation* 110, 452 (2004).
48. Ammar, M., Kassas, N., Chasserot-Golaz, S., Bader, M.-F. & Vitale, N. Lipids in Regulated Exocytosis: What are They Doing? *Frontiers in Endocrinology* 4, 125 (2013).

49. Pfrieger, F.W. & Vitale, N. Cholesterol and the journey of extracellular vesicles. *Journal of Lipid Research* 59, 2255 (2018).
50. Li, J. et al. Integration of lipidomics and transcriptomics unravels aberrant lipid metabolism and defines cholesteryl oleate as potential biomarker of prostate cancer. *Scientific Reports* 6, 20984 (2016).
51. Yue, S. et al. Cholesteryl Ester Accumulation Induced by PTEN Loss and PI3K/AKT Activation Underlies Human Prostate Cancer Aggressiveness. *Cell Metabolism* 19, 393 (2014).
52. Brzozowski, J.S. et al. Lipidomic profiling of extracellular vesicles derived from prostate and prostate cancer cell lines. *Lipids in Health and Disease* 17, 211 (2018).
53. Onozato, M. et al. Amino acid analyses of the exosome-eluted fractions from human serum by HPLC with fluorescence detection. *Practical Laboratory Medicine* 12, e00099 (2018).
54. Kwak, S.S. et al. Autophagy induction by tetrahydrobiopterin deficiency. *Autophagy* 7, 1323 (2011).
55. Skotland, T., Sagini, K., Sandvig, K. & Llorente, A. An emerging focus on lipids in extracellular vesicles. *Advanced Drug Delivery Reviews* 159, 308 (2020).
56. Ramos, I., Malkin, A. & Lyng, F. Current Advances in the Application of Raman Spectroscopy for Molecular Diagnosis of Cervical Cancer. *BioMed Research International* 2015 (2015).
57. Lan, Y.Y., Londoño, D., Bouley, R., Rooney, M.S. & Hacoheh, N. Dnase2a deficiency uncovers lysosomal clearance of damaged nuclear DNA via autophagy. *Cell reports* 9, 180 (2014).
58. Villarroya-Beltri, C. et al. ISGylation controls exosome secretion by promoting lysosomal degradation of MVB proteins. *Nature Communications* 7, 13588 (2016).
59. Takahashi, A. et al. Exosomes maintain cellular homeostasis by excreting harmful DNA from cells. *Nature Communications* 8, 15287 (2017).
60. Krafft, C. Raman Spectroscopy of Proteins and Nucleic Acids: From Amino Acids and Nucleotides to Large Assemblies. *Encyclopedia of Analytical Chemistry*, 1 (2018).
61. Olsztynska, S., Komorowska, M., Vrielynck, L. & Dupuy, N. Vibrational Spectroscopic Study of L-Phenylalanine: Effect of pH. *Applied Spectroscopy* 55, 901 (2001).

62. Zong, S. et al. Facile detection of tumor-derived exosomes using magnetic nanobeads and SERS nanoprobe. *Analytical Methods* 8, 5001 (2016).
63. Kim, J.-H., Kim, E. & Lee, M.Y. Exosomes as diagnostic biomarkers in cancer. *Molecular & Cellular Toxicology* 14, 113 (2018).
64. Lane, R. et al. Cell-derived extracellular vesicles can be used as a biomarker reservoir for glioblastoma tumor subtyping. *Communications Biology* 2, 315 (2019).
65. Prats-Mateu, B. & Gierlinger, N. Tip in–light on: Advantages, challenges, and applications of combining AFM and Raman microscopy on biological samples. *Microscopy Research and Technique* 80, 30 (2017).

Chapter 5. PCR-based strategy for detecting mutations in tomato plants using Surface Enhanced Raman Spectroscopy***

Tomato Yellow Leaf Curl Virus is one of the most devastating diseases threatening the global agricultural industry, and the selection of resistant crops has emerged as the best solution against it. Detection of indels and SNPs conferring this resistance is the quickest way to screen resistant strains, but conventional methods face many disadvantages in carrying this out. In this study, we have developed a PCR-based SERS assay that can relentlessly differentiate between different types of indels and SNPs. This tri-primer assay utilizes mutation-specific forward primers, as well as SERS probes tagged with two types of dyes: FAM and Cy3. These two dyes have unique Raman signature spectra, which helped distinguish between different alleles of an indel in not only the Ty-3 gene but also in the Mi-1 gene, which is responsible for resistance against root-knot nematodes. The same feature of these dyes enabled multiplexing, in which it was possible to detect the indel type of a DNA sample and the zygosity of the allele in a single experiment. This technique successfully differentiated two different SNP-based alleles, providing a promising future in SNP detection as well.



*** This work was in collaboration with Mr. Jayaprakash, Biodecipher, Bengaluru.

5.1. Introduction

Plant viruses have always been a major culprit against the agricultural industry throughout its history. They cause extensive damage to various economically significant crops. Tomato yellow leaf curl virus (TYLCV) is one such type of virus that can cause complete yield loss in the Tomato (*Lycopersicon esculentum*) crop,¹ which is one of the most important culinary vegetables in the world. This disease is transmitted by the Silverleaf whitefly (*Bemisia tabaci*), a major pest for the Solanaceae family.² Characteristic manifestation of TYLCV disease includes extreme hindering of plant height, a decrease in leaf size, curling of leaves, chlorosis on leaves and flowers, and reduction of the production of fruits.³ It has been assessed that around 7 million hectares can encounter TYLCV contamination or mixed viral infections every year.⁴ Plants, in turn, use a variety of defence mechanisms to shield themselves from viral intrusion. RNA silencing is one of the various mechanisms by which non-coding RNAs negatively regulate the expression of viral genes in plants.⁵ RNA silencing in plants utilizes the help of Dicer-like proteins (DCL). A small molecule of RNA is produced (sRNA), which binds with the DCL proteins to form the RNA-induced silencing complex (RISC). This complex is responsible for inhibiting viral protein formation by interacting with the viral messenger RNA (mRNA). However, not all strains of crops can exhibit this phenomenon, which therefore becomes susceptible to infection.⁶

Therefore, screening resistant crops is one of the most effective methods to counter the threats posed by viruses. In the case of TYLCV, plants carrying the genes required for resistance, if challenged by the virus, show low levels of viral replication and spread of infection with moderate or no visual symptoms.⁷ The genes responsible for this are the Ty-1 and Ty-3 genes. These are alleles of the same gene and code for RNA-dependent RNA polymerases from a class of functionally unknown RDR genes, which are thought to be involved with RNA silencing.⁸ Besides TYLCV, infection of root nematodes is another serious problem faced by tomato cultivators worldwide.⁹ They cause decreased yield in adult crops but are lethal to young plants, responsible for approximately 5 % of global crop loss.¹⁰ The Mi-1 gene comes into play during this

viral invasion, as it confers resistance against three common root-knot nematodes, i.e., *Meloidogyne incognita*, *M. javanica*, and *M. arenaria*.¹¹

5.1.1. Mutations – Indels and SNPs

The resistance, discussed above, acquired by some strains of plants is mainly due to mutations like Indels and SNPs. Indel is the accidental introduction or removal of a set of nucleotides from a definite region of the genome, while single nucleotide polymorphism or SNP is a single base-pair substitution at a specific position in the genome (Fig. 5.1).

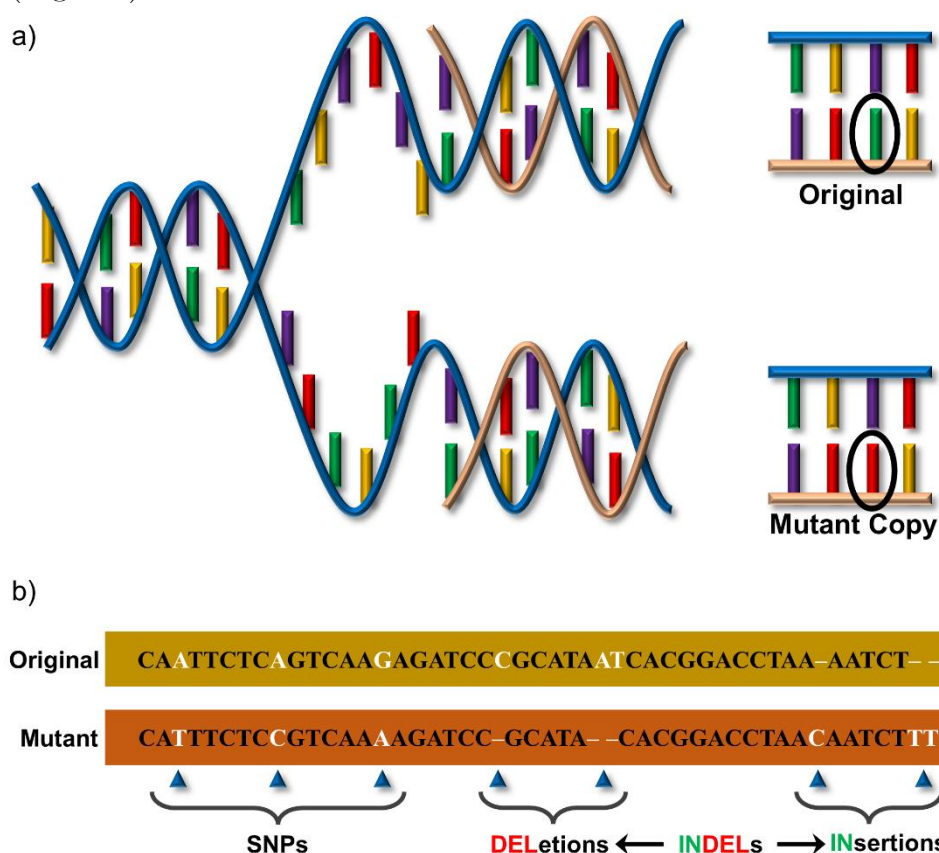


Fig. 5.1. a) Schematic representation of a modification of nucleotide in the DNA sequence – an error during replication, b) representation of the SNPs and the indels mutations.

Both indels and SNPs result in different alleles with distinguishing features like resistance or susceptibility of a plant towards a particular disease. Molecular markers like indels and SNPs for Ty-3 and Mi-1 genes have made screening resistant strains easier. However, conventional methods prevailing for detecting these mutations still have many disadvantages. One of them is direct sequencing, which is undoubtedly highly accurate but expensive and time-consuming. Gel-based methods like RFLP

(Restriction Fragment Length Polymorphism) require a lot of time and a large amount of DNA.¹² Hence there is a requirement to find quick low-cost detections using less quantities of DNA.

5.1.2. PCR based strategies to detect mutations

The solution to the problem of time consumption and requirement of large DNA quantities lies in using PCR-based methods. But they have their disadvantages as well. Methods like DAF (DNA Amplified Fingerprinting) and AP-PCR (Arbitrarily Primed PCR) fail to distinguish between different sequences of similar molecular weight. AFLP (Amplified Fragment Length Polymorphism) cannot differentiate between homozygotes and heterozygotes due to its dependence on dominant markers.¹² Fluorescence-based techniques have circumvented a lot of these issues.¹³ FRET (Förster resonance energy transfer) techniques and molecular beacons have emerged as dependable assays for SNP and indel detection.^{14, 15} But they too are not void of drawbacks, especially when it comes to multiplexing, which implements multiple probes for detecting multiple types of mutations simultaneously. The fluorescence-based techniques use the reporter and quencher relationship for probe activity. Limited availability of these reporter-quencher combinations causes a hindrance in their multiplexing capability. Along with this, a fluorophore's broad fluorescence emission spectrum results in large overlaps and thus reduce the number of dyes that can be used for multiplexing.^{16, 17} Moreover, the detection system is often monochromatic, resulting in non-uniform fluorescence of different dyes, thus restricting the number of probes utilized.

5.1.3. PCR-based Raman spectroscopy to detect mutations

Raman spectroscopy can be considered superior to the above mentioned due to its ability to recognize vibrational signatures of molecules. It relies primarily upon the inelastic scattering of photons, and its intensities can be increased many-fold by using SERS. There is no requirement of a quencher in this method, which also adds to the advantage of this assay being cost-effective compared to the dual labelled probes used

in fluorescence-based PCR methods. SERS assays have been invented to detect specific bacterial DNA sequences,¹⁸ but not many studies have been performed with plant DNA.

5.1.4. Designing oligonucleotide sequences for PCR-based SERS

5.1.4.1. Primer design parameters

Many parameters affect the yield of PCR. The optimal design of primers is essential, and the various parameters that affect them need to be noted.¹⁹

a) Primer Length: Optimal length of the PCR primers is 18-22 bp (base pair). It is long enough to maintain the specificity and short enough for the primer to bind to the template at the annealing temperature.

b) Primer melting temperature (T_m): It is the temperature at which one half of the DNA duplex disassociates to become single-stranded, representing the duplex stability. The GC content of the sequences is a good indication of the primer T_m . The best range of T_m for primer design is 52-58 °C. T_m is calculated using the nearest neighbour thermodynamic theory. The $T_m(K)$ is equal to $\frac{\Delta H}{\Delta S} + R \ln(C)$, where ΔH is obtained by adding up all the dinucleotide pairs enthalpy values of the nearest neighbour base pair and ΔS is the di-nucleotide pairs entropy values of nearest neighbour base pairs, with salt correction.

c) Primer annealing temperature (T_a): Very high T_a produces insufficient primer-template hybridization, hence low PCR product yield. And very low T_a may lead to non-specific products caused by a high number of base pair mismatches. Mismatch tolerance has the strongest influence on PCR specificity. $T_a = 0.3T_m(\text{primer}) + 0.7T_m(\text{product}) - 14.9$, where $T_m(\text{primer})$ and $T_m(\text{product})$ are melting temperatures of the primer and products, respectively.

There are many other parameters that influence the primer design. The GC content should be 40-60 %. The absence of primary, secondary structure in the primer design is essential, as hairpins, self-dimers, or cross-dimers can reduce the PCR yield. Consecutive repeats of the di-nucleotides and long runs of a single base should be avoided. The 3' end should be stable, and the primer shall not be designed on the template's secondary structure to obtain good PCR yields. To have high specificities of the primers, regions of homology must be avoided, i.e., they should amplify other genes in the mixture. Commonly, after designing, the primers are BLASTed to test for specificity.

5.1.4.2. Design parameters for DNA probes

The SERS probes constitute a DNA oligonucleotide sequence with a SERS active dye molecule attached to one end. Uni-molecular or bi-molecular SERS probes can be created for these detection purposes with or without overhangs (Fig. 5.2).²⁰ The composition of the sequence affects the folding of a SERS probe. In the case of the unimolecular probes, the affinity for the target has to be higher than the hairpin formation, which is generally optimised by incorporating mismatches and a reduced number of bases in the hairpin folded region of the SERS probe.²⁰

When no DNA or a non-target DNA is present, the self-complementary region of the SERS probe should occur. The different natures of the uni- and bi- molecular probes result in a difference in their T_m . Unimolecular probes fully form double-stranded DNA than bimolecular ones with higher probability. Commercial production of the bimolecular probes is more cost-effective than the unimolecular probes. Unimolecular probes have been shown to produce higher reproducibility and reduced rate of false positives and negatives than bimolecular probes at lower detection limits.²⁰

Reports suggest that the presence or absence of overhangs in bi-molecular probes does not significantly affect the SERS assay with respect to the sensitivity but has a slightly reduced background. Overall, for direct detection assays, the

unimolecular probes have a performance advantage; else, the results for PCR-based SERS detection can be obtained by both kinds of probes.²⁰

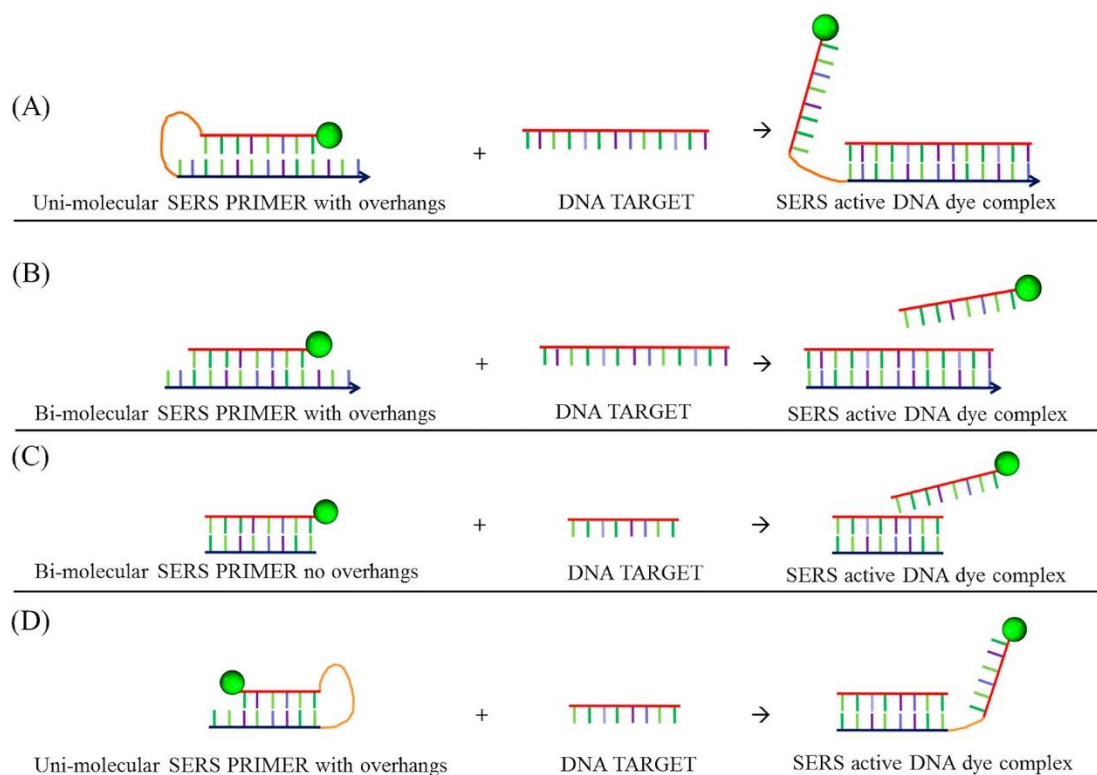


Fig. 5.2. Schematic representation of the various SERS probes: a) unimolecular SERS probe with overhangs, b) bimolecular SERS probe with overhangs, c) bimolecular SERS probe with no overhangs, and d) unimolecular SERS probe with mirrored orientation and overhangs. Copyright © 2013, American Chemical Society

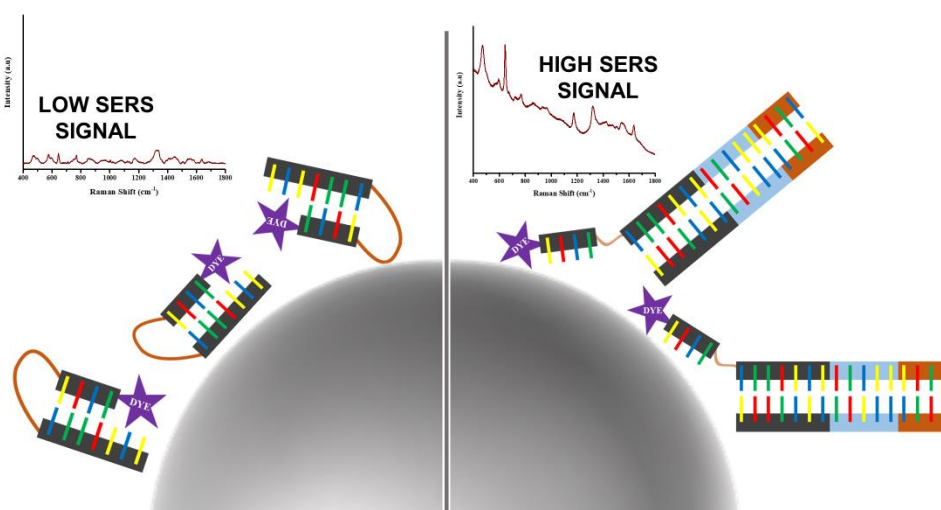


Fig. 5.3. Schematic representation of the functioning of a unimolecular probe.

In the absence of the target DNA, the probe remains folded hence being away from the AgNPs, thus producing low SERS signal. In the presence of the target DNA,

the probe is unfolded, allowing the dye molecule into the vicinity of the AgNPs, resulting in a high SERS signal (Fig. 5.3).

5.1.4.3. Raman active dye molecules for SERS

To conduct SERS, dye molecules with a large Raman scattering cross-section must be chosen. These dyes should be able to easily orient themselves around the AgNPs to produce unique molecularly specific spectra. It is essential that the dyes produce distinct spectra, or at least have a few strong, distinct, non-overlapping peaks that can be used for multiplexing studies. The narrow FWHM of the peaks in a Raman spectrum easily facilitates exploring multiplexing options using a single excitation laser. Due to the presence of the metallic nanoparticles, fluorescence is automatically quenched, allowing for the use of commonly available fluorophores. Some commonly used molecules for such detection include R6G, TAMRA, HEX, Cy3, Cy5, FAM, TET, and ROX, and their SERS spectra are shown in Fig. 5.4.²¹⁻²³

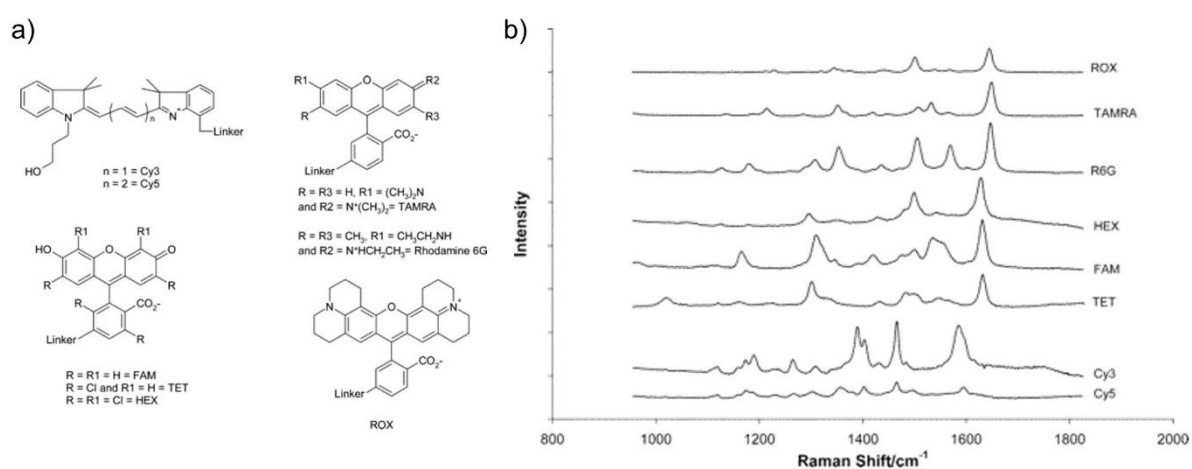


Fig. 5.4. a) Different dye molecules that can be used as SERS probe markers, and b) The SERS spectra corresponding to these dyes. *Copyright © 2004, American Chemical Society*

In this study, we have designed a SERS-based PCR assay capable of screening resistant and susceptible species of Tomato by detecting the discussed mutations involved with Ty-3 and Mi-1 genes and differentiating homozygous and heterozygous alleles among the plant DNA samples. Multiplexing has also been performed, increasing the ease of determining the presence of mutations in unknown DNA sequences.

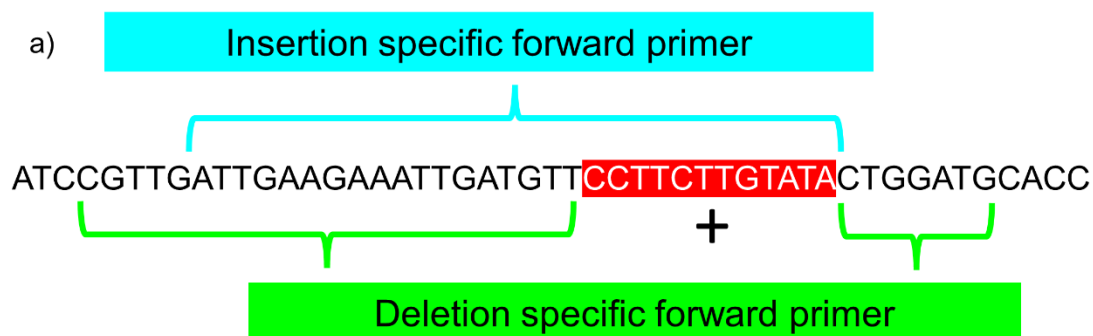
5.2. Materials and Methods

a) Materials:

Spermine tetrahydrochloride (ST) (N,N'-Bis(3-aminopropyl)-1,4-butanediamine tetrahydrochloride), and Silver nitrate (AgNO₃, 99.9999% trace metals basis), sodium citrate were purchased from Sigma Aldrich. Molecular biology grade water was bought from Himedia and TaqMan™ Genotyping Master Mix from Thermo Fisher Scientific. The primers and probes were obtained from Integrated DNA Technologies.

b) Oligonucleotide design:

The genome of *Lycopersicon esculentum* was downloaded from NCBI, and the mutations were located with the help of previously identified markers of Kim et al. in the case of Ty-3 indels. The sequences of the probes and primers were then designed manually.



b)

Uni probe for RESISTANT (INSERTION) allele for INDEL	[FAM]-GTAATCAAAC-[iSp18]-GTTTGATTACACCCCGTAAG
Uni probe for SUSCEPTIBLE (DELETION) allele for INDEL	[Cy3]-TGTTTGC GGA-[iSp18]-TCCGCAAACAGGACAGGTGG
Forward primer corresponding to FAM	GTTTGATTACACCCCGTAAGATTGAAGAAATTGATGTTCTTCTTGTATA
Forward primer corresponding to Cy3	TCCGCAAACAGGACAGGTGGCGTTGATTGAAGAAATTGATGTTCTGGATG

Fig. 5.5. a) The choice of insertion specific (cyan) and deletion specific (green) primers to detect the presence or absence of the indel sequence (red), b) Oligonucleotide sequences of the forward primers corresponding to the insertion specific and deletion specific forward primers, unimolecular probes to detect the tail sequences of the corresponding primers.

In Fig. 5.5a, we have chosen to study the indel sequence (red) in the Ty-3 gene, whose presence or absence decides whether the plant is resistant or susceptible to the

TYLCV. Accordingly, the insertion specific and deletion specific forward primers are chosen, and in Fig. 5.5b, the sequences for the primers and the design of unimolecular probes to detect the tails of the primers is represented as an example.

In designing our strategy, we chose the target of the probe as the complementary region of the tail of the forward primer designed specifically for each indel. This complementary region is amplified during PCR only if the corresponding indel is present in the Tomato DNA sample. The dye tagged unimolecular SERS probe has been designed to have a higher affinity towards the target region than its own complementary strand. Thus, in the presence of a target DNA, self-hybridization is likely to be prevented. This has been brought about by the difference in T_m of at least 15°C in its self-complementary regions.

c) DNA samples:

The tomato seeds were germinated and the leaves were harvested for genomic DNA isolation. when the plants reached the 3rd or 4th leaf stage The DNA samples were extracted from Tomato plants via the CTAB method and supplied by BioDecipher, Bengaluru. The concentration of the DNA was adjusted to $50\text{ ng}/\mu\text{L}$.



Ty-3 – Indel marker analysis (Resistant – 155bp, upper band, Susceptible – 145bp, lower band)

1, 2,332 are tomato plants carrying resistant or susceptible or both genes(heterozygous condition)

For example,

- Plant numbers 6, 20, 27 have resistant gene,
- Plant numbers 1, 2, 5 have susceptible gene and
- Plant numbers 3, 4, 7 have both copies indicating plants are in heterozygous condition.

L – 50 bp DNA ladder

Ty-3 indel marker primers used in this study are:

- *Forward primer* – 5' ATGGGTGATCCGTTGATTGAAG 3'
- *Reverse primer* – 5' CACCAATAGAGCTCAGCCTCC 3'

Fig. 5.6. Gel Data of the indels in the Ty-3 indicating whether the plants have the resistant gene or the susceptible gene.

The zygosity and the resistance/susceptibility of the samples were analysed by BioDecipher via a gel-based method and therefore known to us beforehand. Fig. 5.6 and Fig. 5.7 show the gel data, indicating the resistant and susceptible plants corresponding to the indels in the Ty-3 and Mi-1 plants respectively. *Escherichia coli* plasmid DNA (XL-10 Gold), used as a negative control, was obtained via alkaline lysis protocol.

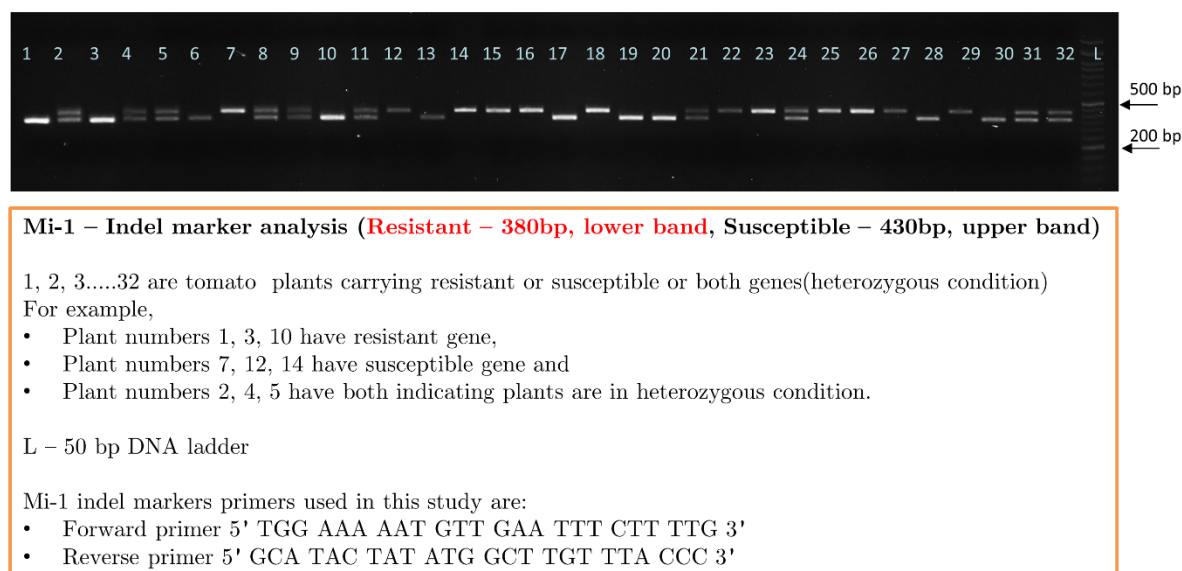


Fig. 5.7. Gel Data of the indels in the Mi-1 indicating whether the plants have the resistant gene or the susceptible gene.

d) Polymerase chain reaction (PCR):

A total volume of 25 μL was used for each reaction. Each reaction tube contained 12.5 μL TaqMan™ Genotyping Master Mix. This Master Mix consists of AmpliTaq Gold™ DNA Polymerase UP (Ultra-Pure), dNTPs, MgCl_2 and other essential components necessary for PCR. 2.5 μL of DNA (50 ng/ μL) was used for each reaction. For single detection, 2 μL of 2.5 μM forward primers and reverse primers each were added, along with 1 μL of 2.5 μM designed SERS primers. For multiplexing studies, 2 μL of both the forward primers and 1 μL of both the SERS primers were added. Volume was made up to 25 μL using Molecular Biology Grade water. The PCR reactions were carried out in T100 Thermal Cycler (Bio-Rad Laboratories, Inc). As seen in Fig. 5.8, the cycling protocol was as follows: 15 minutes at 94°C for hot start activation, thenceforth, 30 repeats of 45 s at 94 °C, 60 s at 62 °C and 60 s at 72 °C,

followed by a final extension at 72 °C for 1 minute. After the process was completed, the samples were cooled down gradually to room temperature.

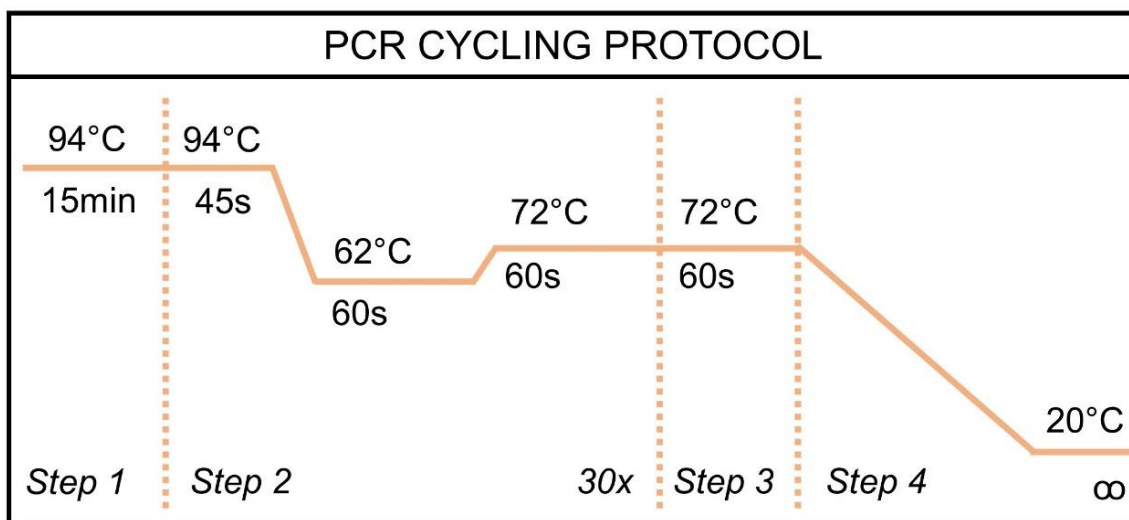


Fig. 5.8. Schematic indicating the PCR cycling protocol.

e) Sample preparation and SERS experimental parameters:

After the PCR was complete, 2.5 μL was taken from each reaction tube and added to 125 μL of 1X Phosphate Buffer Saline (137 mM NaCl, 2.7 mM KCl, 4.3 mM Na_2HPO_4 , 1.47 mM KH_2PO_4). Then, 10 μL of 400 μM (or 0.01 M, depending upon the experiment) Spermine Tetrahydrochloride (ST) was added to the mixture, and the solution was incubated for 6 minutes.

The synthesised AgNPs were concentrated to 20X at 7200 rpm for 10 min. 8 μL of the concentrated AgNPs was added to 2 μL of the incubated PCR product. The Raman spectra of this mixture were collected in liquid mode using LabRam HR Evolution. 532 nm incident laser at 2.3 mW power, 1800 grooves/mm grating and acquisition of 10 s per window was used to collect the SERS spectra. The obtained Raman spectra were baseline corrected using a user-defined baseline. The required Raman peaks were fit using a Lorentzian curve, and the statistical distribution of intensities was mapped using box plots.

5.3. Results and Discussions

5.3.1. PCR-based SERS strategy

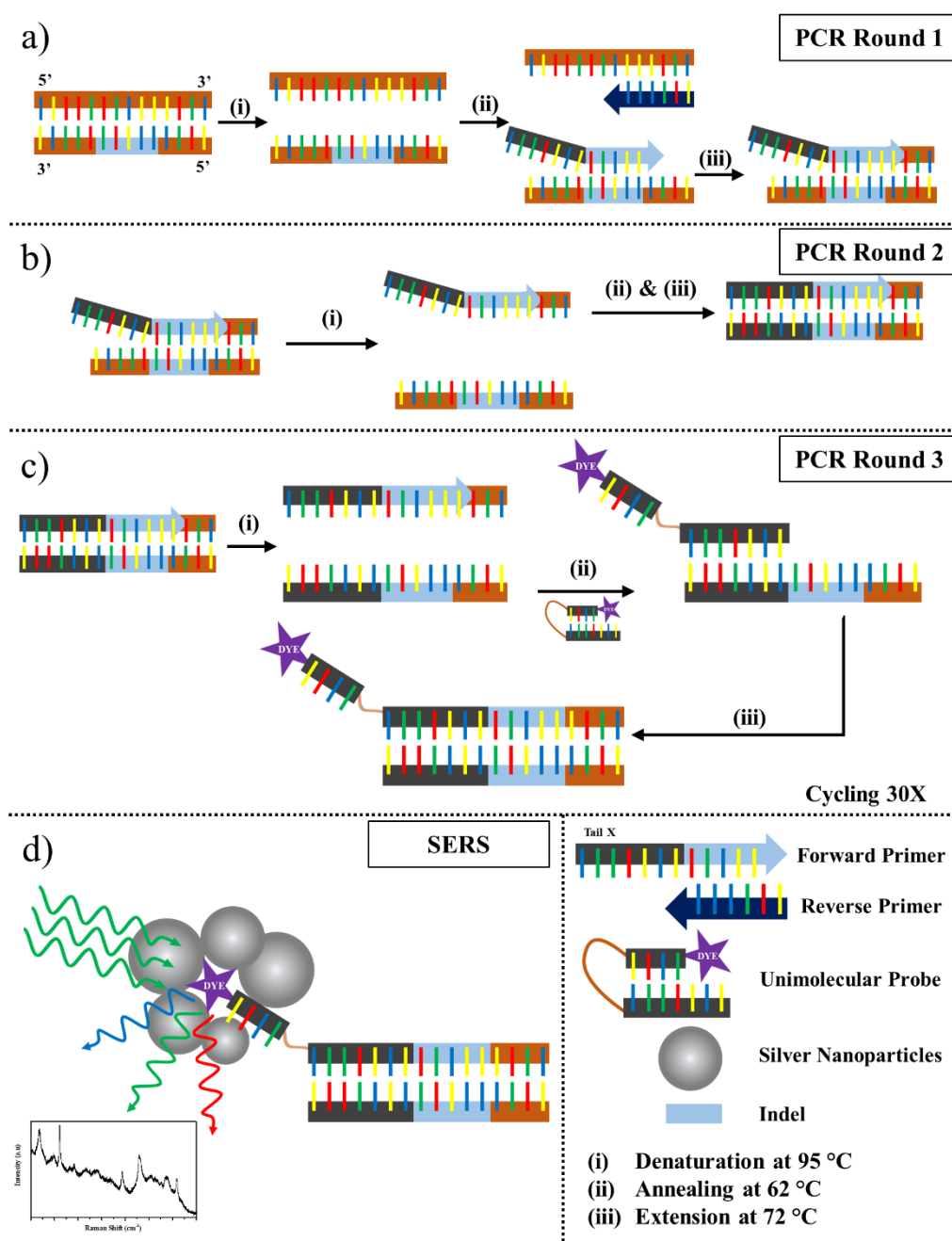


Fig. 5.9. Schematic representation of the PCR-based SERS strategy. a) Annealing of mutation-specific forward primer and reverse primer, with the non-complementary tail of the forward primer failing to anneal. b) Formation of the unique, complementary region of the hanging tail. c) SERS probe binds to its target, the complementary region of the tail. d) Increased Raman signal from the exposed dye tagged region of the SERS probe.

The main principle of our strategy lies in the fact that the DNA binds well to a complementary region in the other strand as compared to a random sequence. So,

our strategy involved the design of a unimolecular probe, in which the dye molecule remains buried when the probe is closed. When the target DNA is present, our strategy allows for an increased selective interaction with the probe, thus freeing the dye molecule of the probe to allow for an enhanced Raman signal. While, in the absence of the target, the designed unimolecular probe self hybridizes, thus giving a weaker Raman signal.

As seen in Fig. 5.9, the steps involved in the PCR followed by SERS is represented schematically. After the denaturation step of PCR, the allele-specific forward primer and the reverse primer hybridises to their specific targets during the annealing phase. The tail of the forward primer fails to bind to the DNA, but the other part of it gets extended in the next step. In the second round of PCR, the reverse primer extends to create the complementary region of the tail. In the third round of the PCR, the SERS probe binds with the tail's complementary region allowing the dye tagged part to become free. So, once the probe binds to the target, the dye becomes accessible to the silver nanoparticles during SERS.

We chose two dyes with minimum peak overlaps in their spectra for our studies: FAM (6-Carboxyfluorescein) and Cy3 (Cyanine-3). We designed the unimolecular probes to have a FAM-tag for detecting the insertion indel DNA and the Cy3-tag for the deletion indel. As seen in Fig. 5.10, the FAM and Cy3 molecules have distinct Raman spectral signatures. The 1636 cm^{-1} peak originating from the Xanthene ring's C-C stretching²⁴ is a unique spectral feature of the FAM spectra, hence was selected for the analysis of FAM tagged oligo containing experiments. Similarly, for the Cy3 tagged oligos, the 1589 cm^{-1} peak corresponding to the C=N stretch was used for the analysis.²⁵

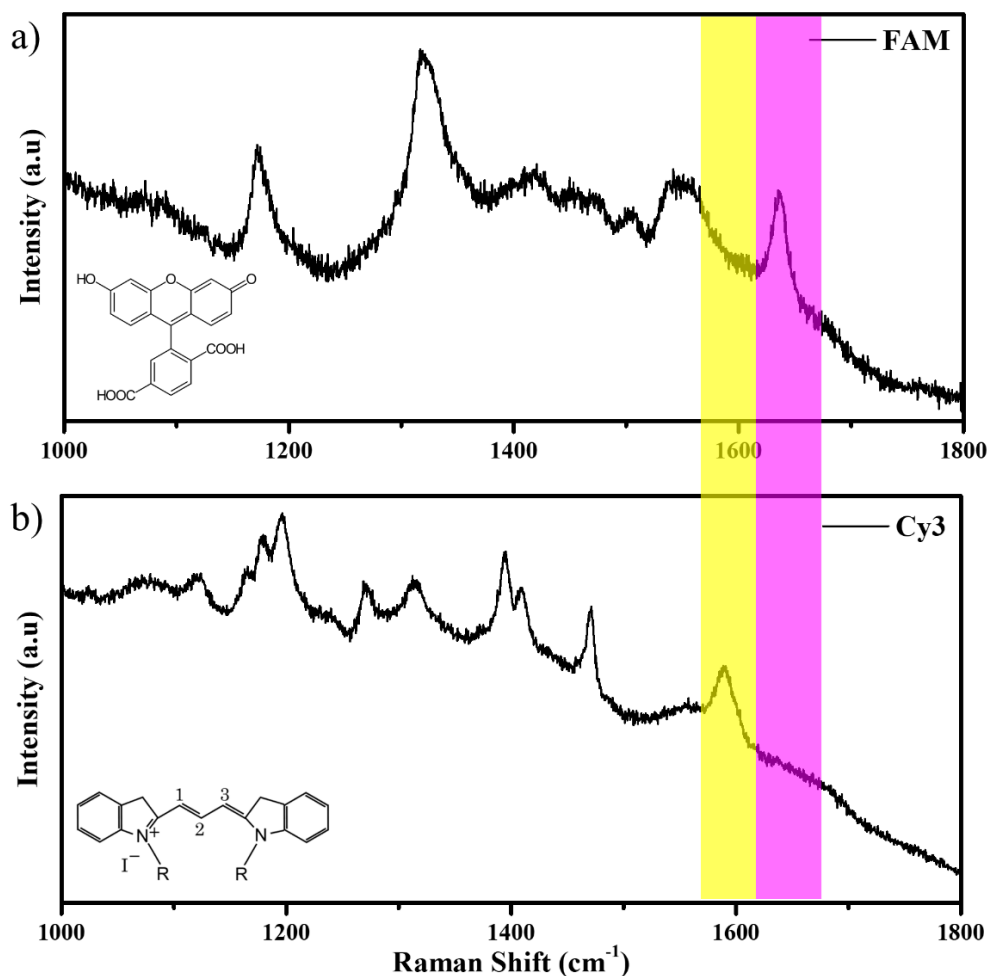


Fig. 5.10. Raman peaks in the SERS spectra corresponding to the a) FAM and the b) Cy3 dyes. The yellow highlighted region with a peak at 1589 cm⁻¹ represents the unique peak corresponding to Cy3, and the pink highlighted region with a peak at 1636 cm⁻¹ represents the unique peak corresponding to FAM.

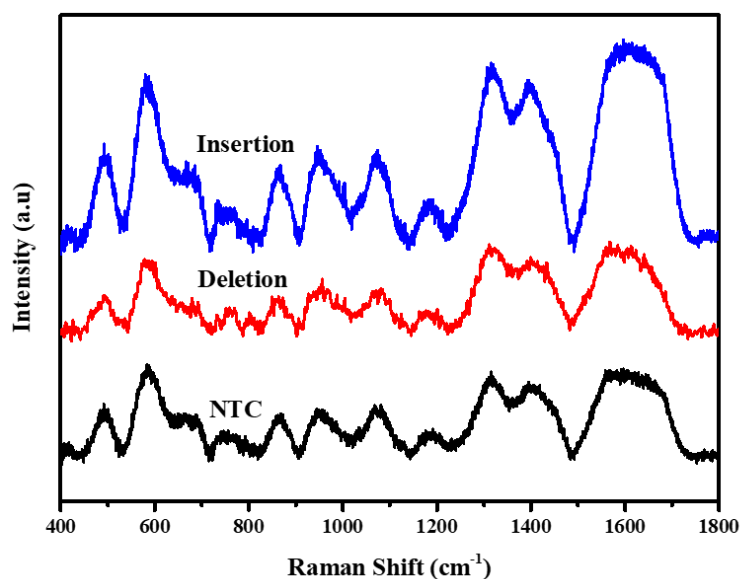


Fig. 5.11. Raman spectra of a mixture of AgNPs with the PCR products, i) containing no DNA (NTC) (black), ii) deletion indel DNA (red) and iii) insertion indel DNA (blue) using the FAM-tag unimolecular probe.

It is evident from Fig. 5.11 that when the FAM-tag unimolecular probe is used, the insertion specific DNA that was supposed to give a SERS spectrum with the peaks corresponding to the FAM dye did not. There was no signal of the dye as both the DNA backbone and the silver nanoparticles are negatively charged, hence repelling each other. Spermine tetrahydrochloride (ST) neutralizes the charge of the DNA backbone,^{26, 27} thus allowing for the nanoparticles to come closer to the dye and generating an increased SERS signal. ST also aggregates nanoparticles, thus adding to our advantage of increasing the formation of SERS hotspots for an improved signal.^{28, 29} Our study was carried out using two different concentrations of ST, 400 μ M and 0.01 M. These two concentrations were the two extreme values provided by Graham et al. that could properly differentiate the presence and absence of a sequence through SERS.

5.3.2. Detection of Ty-3 indel by PCR-based SERS

The oligo sequences corresponding to the forward and reverse primers, insertion and deletion specific unimolecular probes designed to detect the presence or absence of an indel in the Ty-3 gene, are given in Table. 5.1. These were used to conduct the PCR experiments, and then the SERS studies were conducted on the obtained PCR products.

Type	Sequence
Forward primer specific to Insertion	GTTTGATTACACCCCGTAAGATTGAAGAAATTGATGTTCCCTTCTGTATA
Forward primer specific to Deletion	TCCGCAAACAGGACAGGTGGCGTTGATTGAAGAAATTGATGTTCTGGATG
Reverse primer	GTTTTTGCCCCTGCTCCTTG
Probe for detection of tail of insertion specific forward primer	[6FAM]-TGCCCTTGTA-[iSp18]-TACAAGGGCACGTAGTTCAC
Probe for detection of tail of deletion specific forward primer	[Cy3]-GCTCACCCAC-[iSp18]-GTGGGTGAGCCTGTTACACC

Table. 5.1. Oligo sequences for Ty-3 indel detection.

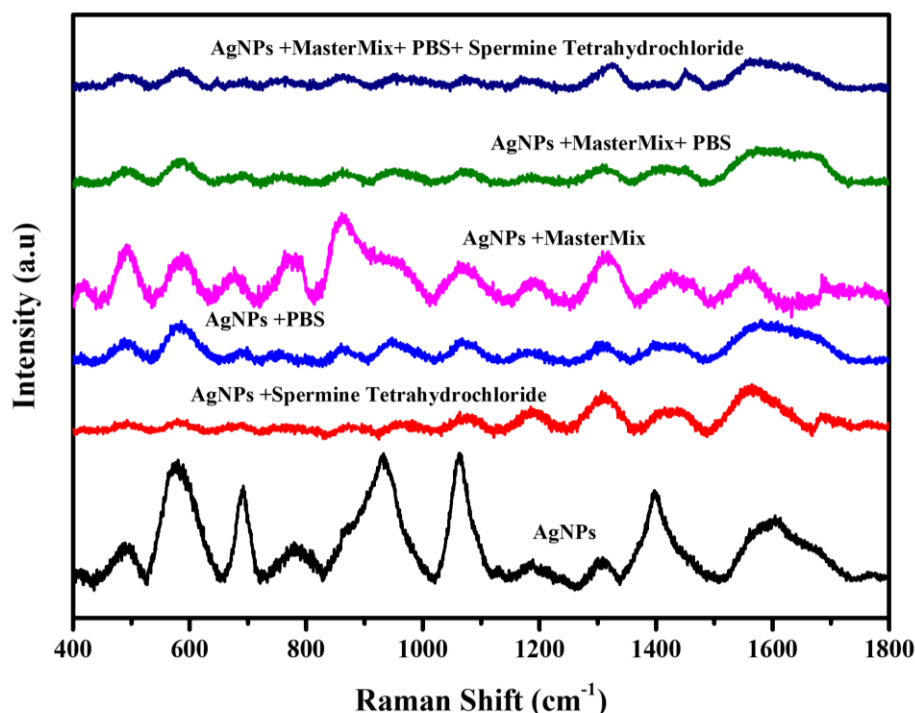


Fig. 5.12. Raman spectra of all the controls including AgNPs, mixtures of AgNPs with spermine tetrahydrochloride, PBS, Mastermix, Mastermix+PBS, and Mastermix+PBS+ spermine tetrahydrochloride.

Before conducting the SERs experiments, the Raman spectra of the AgNPs, AgNPs in PBS, AgNPs with ST, AgNPs with mastermix, AgNPs with mastermix and PBS, and AgNPs with Mastermix, PBS and ST were obtained as controls to be sure of no background overlapping peaks with the spectral peaks of the dyes (Fig. 5.12). All the SERS experiments are a consolidated picture of many spectra collected over 3 different studies. These results proved consistent over the experimental repeats, and their averages shall be used to represent the overall result. For the FAM-tag probe mixtures, we obtained the intensity of the 1636 cm⁻¹ peak through a Lorentzian fit after a user-defined baseline correction, and similarly, the intensities of the 1589 cm⁻¹ peak were obtained for the Cy3-tag probe mixtures. The no-template control (NTC) sample was prepared by replacing the DNA with water in the PCR tube to act as a control for the SERS experiments.

Using the obtained PCR product and 400 μ M ST, the SERS experiments were conducted. As seen in Fig. 5.13a, for the FAM-tag probes, the box plot of the intensities of the 1636 cm⁻¹ peak shows an average higher intensity for the insertion specific DNA,

while the average was lower for the NTC and deletion specific DNA. The Raman spectrum obtained by averaging over the multiple runs is represented in Fig. 5.13b, where the peak intensity difference is clearly visible. In the presence of the insertion sequence (Homozygous resistant crop) in the Ty-3 gene, the tail corresponding to the insertion specific forward primer had its complement synthesized more in amount compared to the deletion specific forward primer during the PCR. As a result, the FAM tagged probe hybridized to the complementary region, freeing the dye and thus generating an increased SERS signal (Blue). In the case of NTC and negative samples, the absence of the target sequence (deletion) forced the FAM tagged probe to self-hybridize, thus yielding a lower Raman intensity.

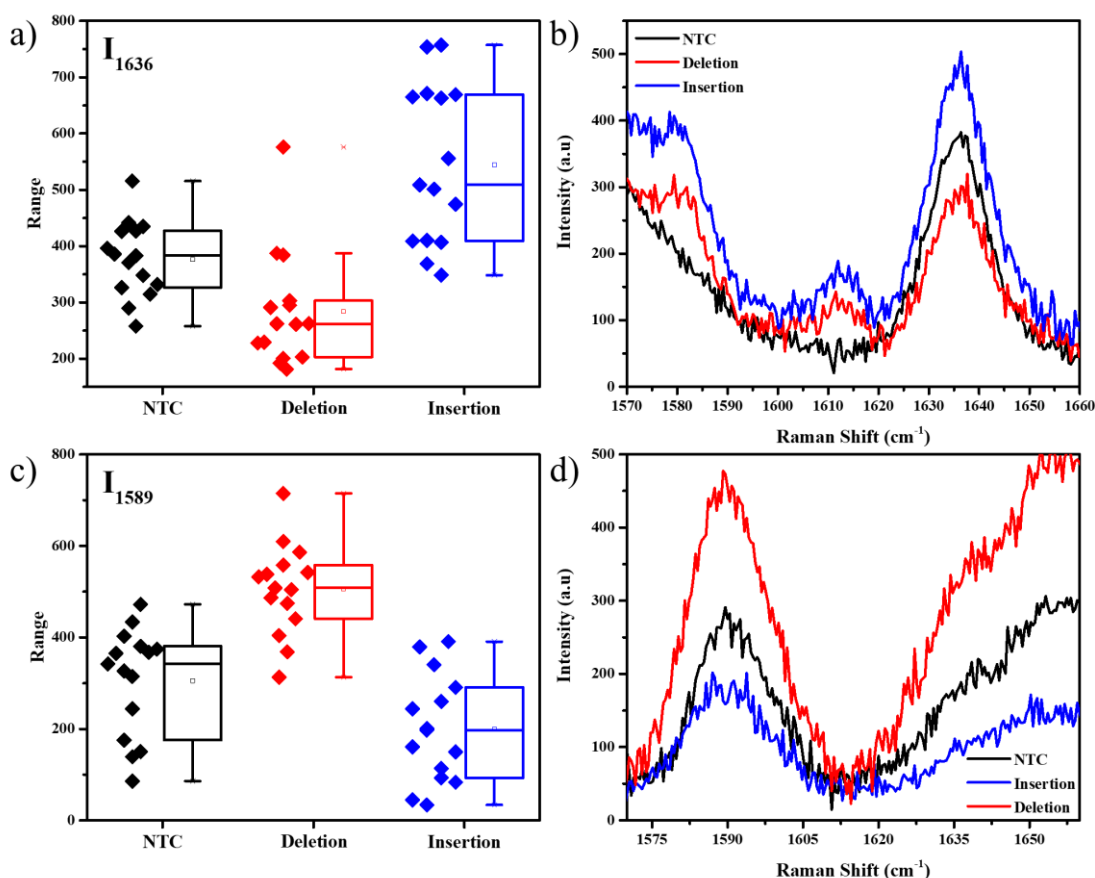


Fig. 5.13. 1636 cm⁻¹ peak analysis for the FAM-labelled SERS primer containing PCR products, a) Box plot representing the 1636 cm⁻¹ peak intensities, b) SERS spectra in the unique regions. 1589 cm⁻¹ peak analysis for the Cy3-labelled SERS primer containing PCR products, c) Box plot representing the 1589 cm⁻¹ peak intensities, d) SERS spectra in the unique regions. Black – no DNA-Template control (NTC), red – Susceptible DNA (deletion indel), and blue – Resistant DNA (insertion indel).

Similarly, for the Cy3-tag probes, the box plot of intensities of the 1589 cm^{-1} peak (Fig. 5.13c) and the average Raman spectra (Fig. 5.13d) represented a higher average intensity for the deletion specific DNA as compared to the NTC and the insertion specific DNA. The complementary region of the tail of the deletion specific forward primer is the target of the Cy3-tagged probes. This region was amplified when the DNA template was that of a Susceptible Tomato plant. From Fig. 5.13, the differences in the box plots clearly show that our strategy can be used to differentiate between susceptible and resistant species of the tomato plant based on their Ty-3 gene composition. In the NTC sample, the DNA is absent, and so the spermine tetrahydrochloride (ST) only affects the existing unimolecular probes and hence resulting in a quite uncertain behavior of the sample. In the negative DNA sample, there is a competition of the DNA with the probe oligos to come closer to the AgNPs and hence an overall lower count of the SERS intensity as compared to the NTC where the unimolecular probes do not face any competition.

5.3.3. Multiplexing studies using PCR-based SERS

The unique isolated peaks of FAM and Cy3 make SERS multiplexing a possibility, thus helping determine the nature of an unknown Tomato DNA sample. Both the probes and their corresponding forward primers were added to the PCR tubes for multiplexing studies. The peak intensity ratio I_{1589}/I_{1636} , corresponding to the signature peaks of Cy3 and FAM, were obtained for all the spectra. A box plot for these intensity ratios is represented in Fig. 5.14a. The Fig. 5.14b represents the corresponding peaks in the SERS spectra of the various samples. In the absence of any DNA, both the dye molecules would be equally but limitedly accessible to the AgNPs. As their Raman scattering cross-sections would be different, slight differences in the intensities corresponding to their peaks are observed. The average intensity ratio of the NTC sample was obtained from the box plot as 1.2. This value was used as a reference for further analysis. A higher ratio compared to NTC indicated a larger number of Cy3 dye molecules accessible to the AgNPs than the FAM, thus confirming that the probes were in the presence of a DNA from the susceptible type Tomato gene,

i.e. the deletion indel. A higher intensity ratio than NTC indicates the DNA from a susceptible Tomato gene, while a lower ratio indicates the DNA obtained from a resistant type of Tomato gene.

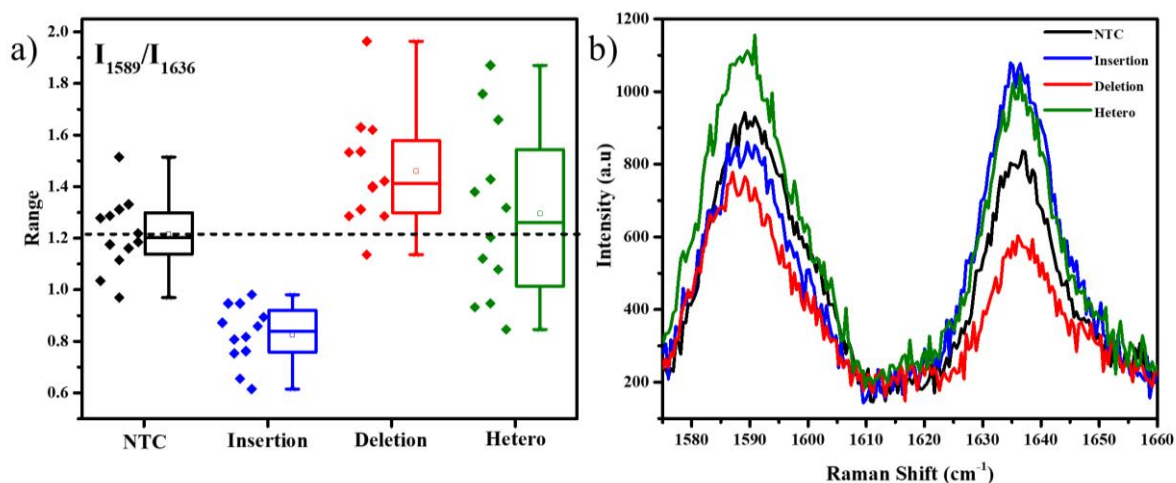


Fig. 5.14. Multiplexing SERS spectral analysis of the PCR products containing both the FAM-tagged and Cy3-tagged probes – Intensity ratio of 1589 cm⁻¹ and 1636 cm⁻¹ peaks, a) box plot representation and b) SERS spectra of the unique spectral region. Black – no DNA-template control, blue – Resistant DNA (insertion indel), red – susceptible DNA (deletion indel), green – Heterozygous DNA (both deletion and insertion indel).

This multiplexing experiment was conducted for a random DNA. Here we chose DNA extracted from *E.coli*. An absence of any of the two probe-specific sequences allowed for a few non-specific bindings equally with both the probes, thus eventually resulting in an intensity ratio equal to NTC (Fig. 5.15), thus confirming the absence of either of the 2 target DNA sequences.

Many of the DNA samples extracted from the plants are heterozygous in nature, and it becomes important to study how these samples behave in these experimental conditions. Multiplexing study was conducted for heterozygous DNA samples (Fig. 5.14), containing insertion and deletion indels. As seen in Fig. 5.14a, the box plot appeared widely scattered due to signals from both the alleles. However, the mean value at 1.3 was close to NTC, as the signals from both FAM and Cy3 had identical intensities. This can be explained by the heterozygous DNA's nature of having an equal number of insertion and deletions indels. Hence, the overall result is an average equivalent to NTC, or *E.coli*, where either no DNA is present, or the DNA doesn't match.

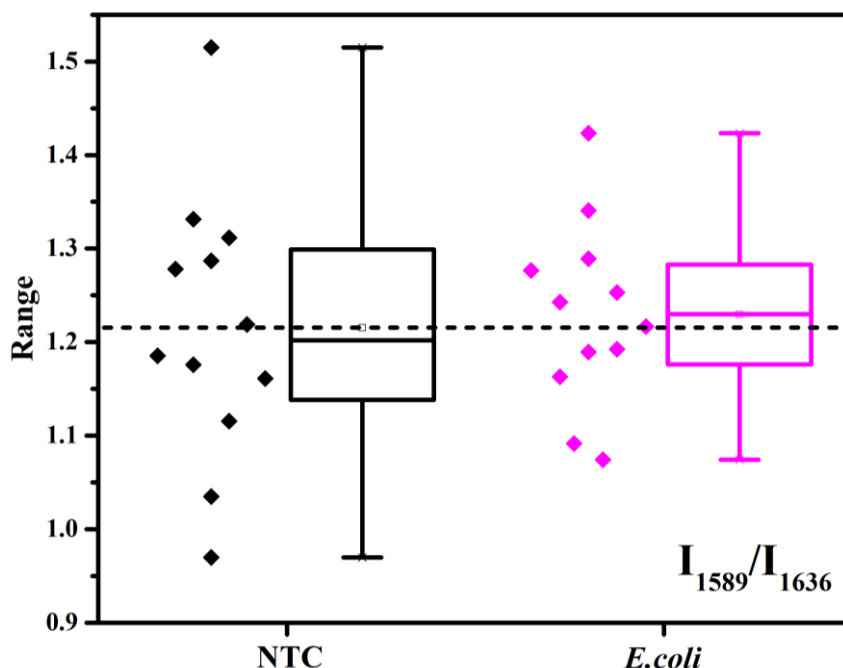


Fig. 5.15. Box plot corresponding to multiplexing studies of the *E.coli* DNA.

5.3.4. Detecting Mi-1 indel using PCR-based SERS

We used the same protocol to evaluate the assay's dependability in another situation by detecting an indel in a different gene involved with a separate disease. Mi-1 gene present in Tomato is involved in conferring resistance against root-knot nematodes. Deleting a specific sequence causes the plant's resistance, while insertion causes it to become susceptible to the nematode's attack. Markers for these mutations were used to develop the primers and probes accordingly (Table. 5.2)

Type	Sequence
Forward primer specific to Insertion	CAACCTGGCGGGTATATTCAGGACGTGAACTGTCACATTTC
Forward primer specific to Deletion	GAGGGTTAATCCCTCGGACTCAATTTGTTTGACACTTTCTCCGC
Reverse primer	GTTTCTAAAATTTTGGAAATATTCTGGC
Probe for detection of tail of insertion specific forward primer	[6FAM]-CGCCAGGTTG-[iSp18]-CAACCTGGCGGGTATATTCA
Probe for detection of tail of deletion specific forward primer	[Cy3]-ATTAACCCTC-[iSp18]-GAGGGTTAATCCCTCGGACT

Table. 5.2. Oligo sequences for Mi-1 indel detection.

We maintained the same system of using FAM-tagged probes to detect the insertion and Cy3-tagged probes to detect the deletion. Results similar to the previous experiments were obtained, showing distinct differences between the intensities of the dyes in the presence of the two different mutations. (Fig. 5.16a and Fig. 5.16b) This experiment proves that the strategy that we have developed can be used to detect indels in any DNA system.

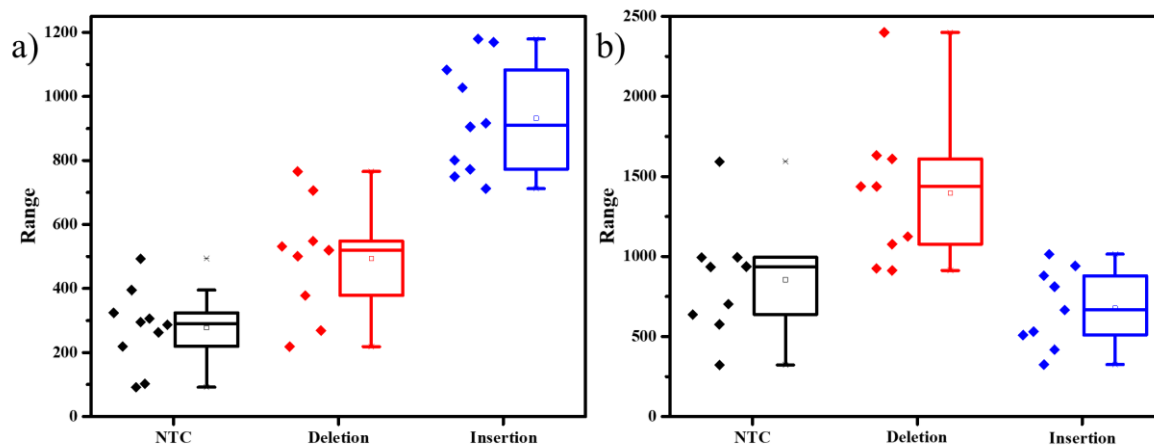


Fig. 5.16. Box plots of SERS spectral analysis of the PCR products for Mi-1 gene containing a) Intensity of 1636 cm⁻¹ peak corresponding to FAM-tagged probe and b) Intensity of 1589 cm⁻¹ peak corresponding to Cy3-tagged probes. Black – no DNA-template control, red – Resistant DNA (deletion indel), blue – Susceptible DNA (insertion indel).

5.3.5. Using higher concentration ST in PCR-based SERS

The SERS experiments with 400 μ M ST were able to differentiate the insertion indels from the deletion indels clearly on an average but had a few data points in the box plots that were overlapping. This indicates a comparable number of closed unimolecular probes to the opened unimolecular probes. To increase the access of the dye molecules in the opened probes to the AgNPs, we further increased the ST concentration to a higher value of 0.01 M. As seen in the TEM figures, represented in Fig. 5.17, with an increase in ST concentration, larger clusters of AgNPs are formed. The inset in Fig. 5.17a shows a slight grey tinge to the AgNPs, and the TEM image in Fig. 5.17a shows agglomerated AgNPs in size range of 500-800 nm, for 400 μ M ST. For an ST concentration of 0.01 M, the inset of Fig. 5.17b shows a dark grey colour to the AgNPs, indicating the formation of larger particles, which is very clearly seen in the TEM image (Fig. 5.17b), indicating agglomerates in size range of 2-3 μ m. The

formation of such large aggregates allowed for more SERS hotspots for the molecules to sit amidst.

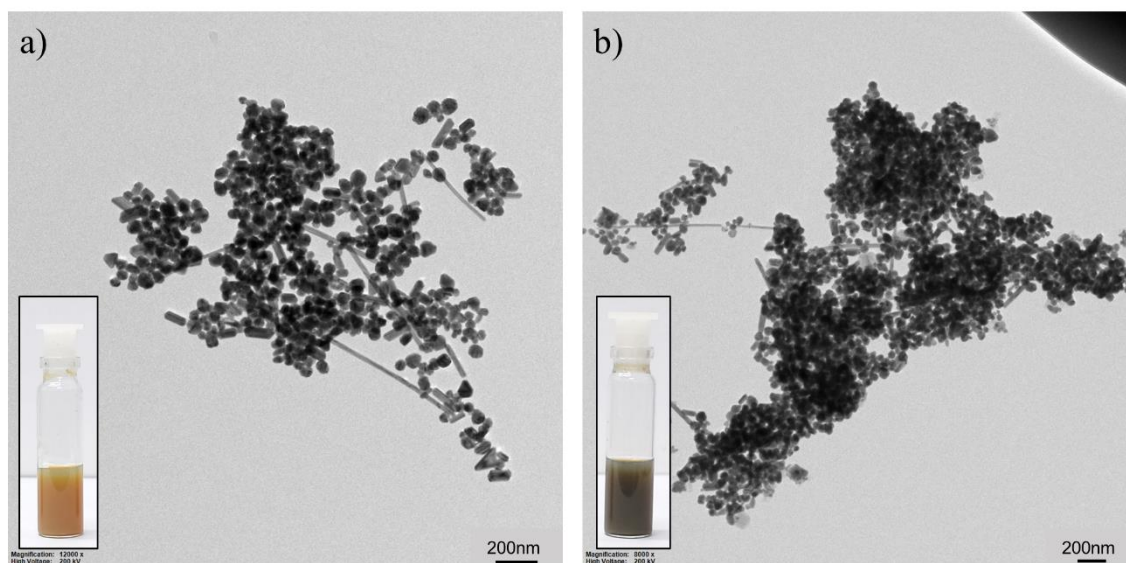


Fig. 5.17. TEM images of AgNPs with a) 400 μM ST indicating aggregates in size range of 500–800 nm, and b) 0.01 M ST indicating aggregates of 2–3 μm size. Inset – optical image of the mixtures show, a) slight tinge of grey colour of mixture, b) dark grey colour of the mixture with 0.01 M ST.

SERS studies on the PCR products with 0.01 M ST indicated a larger separation in the intensities represented by the box plots. As seen in Fig. 5.18a, the average SERS intensity of the FAM peak corresponding to the insertion specific DNA was considerably large compared to the NTC and deletion specific DNA. Similarly, in Fig. 5.18b, the SERS intensity of the Cy3 peak was considerably higher for the deletion specific indel compared to NTC or the insertion specific indel. In, Fig. 5.18, the box plots of the heterozygous samples marked their corresponding average values midway to the two homozygous samples. This was as expected, as the heterozygous sample is a mix of both. Similarly, the average intensities of the *E.coli* DNA matched the NTC and deletion (for FAM-tagged probe) or insertion (for Cy3-tagged probe), as there was no amplification of this DNA. Therefore, a higher concentration of ST allowed for better charge neutralising and more SERS hotspots, hence easy differentiation of the indels under study. From the box plots in Fig. 5.18a, the plant sample with the insertion indel can be distinguished from that with deletion indel with $> 90\text{--}95\%$ accuracy.

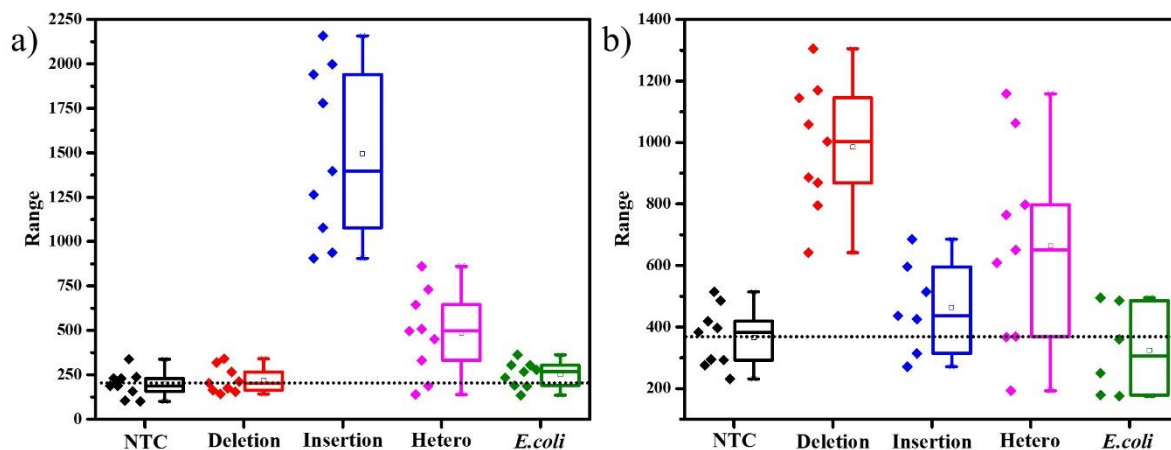


Fig. 5.18. Box plot representations of intensities of unique peaks from SERS analysis, a) 1636 cm^{-1} peak for the FAM-tagged SERS primer containing PCR products, b) 1589 cm^{-1} peak for the Cy3-tagged SERS primer containing PCR products. Black – no DNA-Template control (NTC), red – Susceptible DNA (deletion indel), and blue – Resistant DNA (insertion indel), magenta – Heterozygous DNA (both deletion and insertion indels), green – *E.coli* (negative control – bacterial DNA).

We conducted multiplexing studies using 0.01M ST for the SERS studies and made box plots using the ratios of the intensities of the Cy3-tag peak, 1589 cm^{-1} and that of the FAM-tag peak, 1636 cm^{-1} . We clearly distinguished the resistant and the susceptible DNA using this multiplexing technique, as seen in Fig. 5.19.

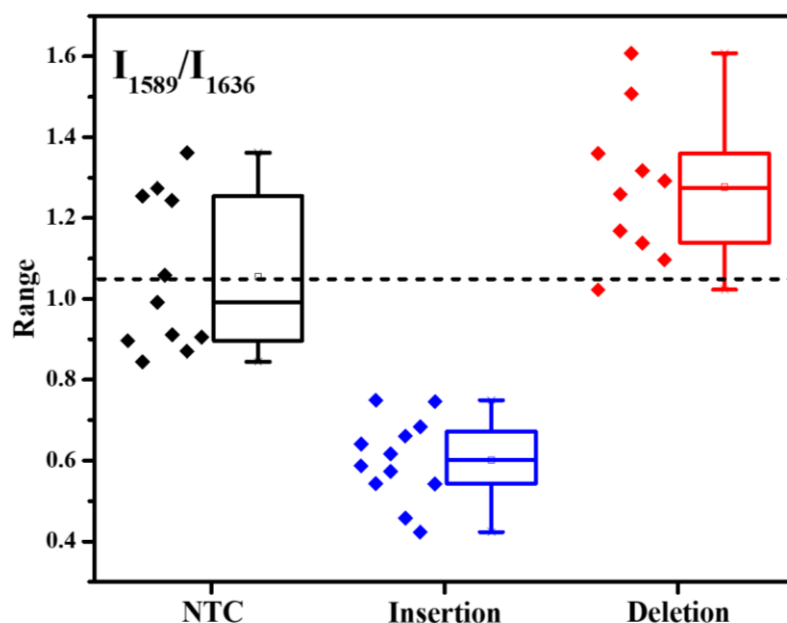


Fig. 5.19. Box plot for multiplexing SERS spectral analysis of the PCR products containing both FAM-tagged and Cy3-tagged probes using 0.01M ST – Intensity ratio of 1589 cm^{-1} and 1636 cm^{-1} peaks, blue – Resistant DNA (insertion indel), red – susceptible DNA (deletion indel).

5.3.6. Detecting SNP in Ty-3 gene using PCR-based SERS

The Ty-3 gene also exhibits the phenomenon of Single Nucleotide Polymorphism, in which the C-allele confers the resistance against TYLCV, but the T-allele makes it susceptible.⁸ We developed probes based on the available markers, where the FAM-tag probe was designed to detect the C-allele.

Type	Sequence
Forward primer specific to Insertion	CAACCTGGCGGGTATATTCAGGACGTGAACTGTCACATTTTC
Forward primer specific to Deletion	GAGGGTTAATCCCTCGGACTCAATTTGTTTGACACTTTCTCCGC
Reverse primer	GTTTCTAAAATTTTGGAAATATTCTGGC
Probe for detection of tail of insertion specific forward primer	[6FAM]-CGCCAGGTTG-[iSp18]-CAACCTGGCGGGTATATTTCA
Probe for detection of tail of deletion specific forward primer	[Cy3]-ATTAACCCTC-[iSp18]-GAGGGTTAATCCCTCGGACT

Table. 5.3. Oligo sequences for Ty-3 SNP detection.

As expected, the mean intensity of the 1636 cm^{-1} peak was the highest in the case of C-allele, while T-allele showed an average lower intensity, even lesser than NTC itself (Fig. 5.20a and Fig. 5.20b). Thus, it is evident that this assay can also detect SNP in plant genomes. As this strategy involves the detection of the tail of the forward primer and not the original DNA strand itself, over the course of the PCR cycles, the probability of non-specific binding reduces. Also, as a part of the design of the forward primer for SNP detection, we have created a mismatch in the 3rd base pair; thus, if the required allele is absent, there is higher confidence of reduction in non-specific hybridisation. These assure that our strategy can be more dependable for SNP detection. With the present set of data points, we can predict an accuracy rate of about 80% for the differentiation of the C and the T SNP in the DNA of a random sample. With automation, there is a reduction in the manual errors and larger samples that can be obtained and hence the accuracy can be further increased. Still, we have

demonstrated that the present strategy is able to establish a trend of differentiating SNPs and indels.

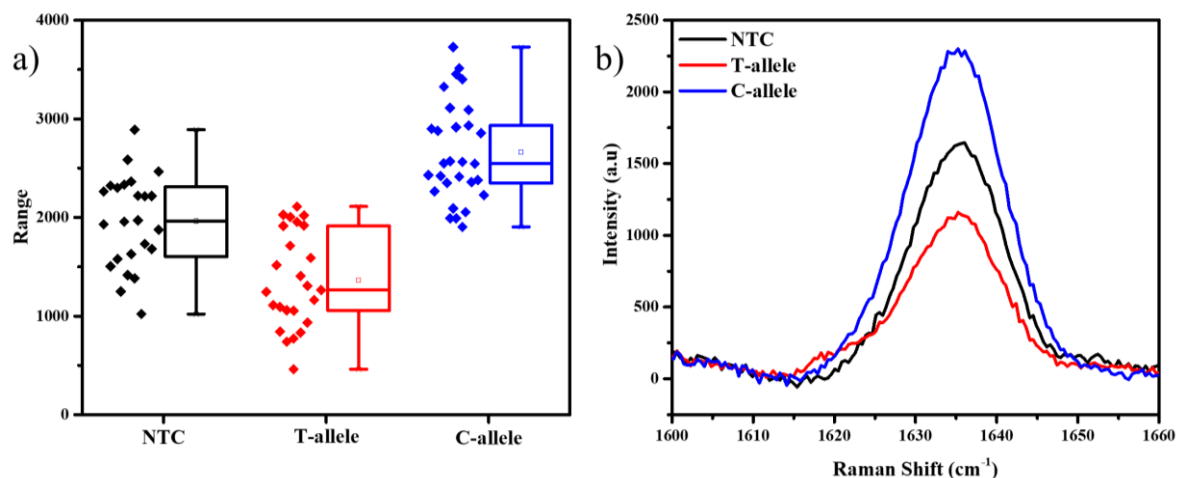


Fig. 5.20. 1636 cm^{-1} peak analysis for the FAM-labelled SERS primer containing PCR products, a) Box plot representing the 1636 cm^{-1} peak intensities, b) SERS spectra in the unique regions. Black – no DNA-Template control (NTC), red – Susceptible DNA (T-allele), and blue – Resistant DNA (C-allele).

Techniques:	Gel-based methods (RAPD, RFLP)	Sequencing methods (Sanger, pyrosequencing, Mini sequencing)	TaqMan TM	Fluorescent molecular beacon	PCR-based SERS detection
Features					
Detection	Endpoint	Real-time/Endpoint	Real-time	Real-time	Endpoint
Detection Platform	Fluorescence or Radioactive	Fluorescence/Luminometric etc.	Fluorescence	Fluorescence	Raman
Specialized Equipment required	Electrophoresis chamber	Colorimeter/Pyro sequencer/MALDI-TOF/Flow cytometer etc.	Fluorescence monitoring thermocycler	Fluorescence monitoring thermocycler	Normal thermocycler, Raman Spectrometer
Multiplexing Ability	Moderate	-	Moderate	Moderate	High
Cost	Low	High	High	High	Moderate
Specificity	Low	High	Moderate	Moderate	Moderate
Time	Time-consuming	Time-consuming	Time saving	Time saving	Time saving

Table. 5.4. Comparison of the existing techniques with our PCR-based SERS detection assay.

As shown in the Table. 5.4 (inspired from Lakhanpaul, T.J. et al³⁰), we compared the existing techniques with the PCR-based SERS detection assay. In brief,

the PCR-based SERS technique is comparable to the many other techniques, with a high ability to multiplex, quick scans and low to moderate operational costs.

5.4. Conclusions

It is pretty evident from the above observations that our PCR-based SERS assay can easily distinguish between the presence and absence of a specific sequence in multiple genes, distinctly resulting in the detection of specific indels in the tomato genome. It can also differentiate between homozygous and heterozygous alleles. Moreover, we have also shown that using different dyes, it is possible to perform multiplexing to study the indel-based resistance in the crop, which makes it possible to learn the same about unknown DNA samples in a short period. In addition to that, the assay has also been able to differentiate between two types of alleles showing SNP. Given the ability of Raman spectroscopy to read the molecular fingerprints of the dyes being used, this technique can be accommodated to allow the usage of more than two dyes and thus be fine-tuned to do multiplexing to search for multiple SNPs in a single experiment. The simplicity of this assay also leaves space for automated detection by machines using well plates, which reduces not only the manual errors but also the overall time of the assay. Without loss of generality, through our work, we have demonstrated that without automation and large data sets also, we can establish the trend and differentiate various alleles. Therefore, this technique has the potential to be a highly dependable detection tool for indels and SNPs in the agricultural industry; not only for diseases like TYLCV and root-knot nematodes but also for the selection of desirable lines and thereby promoting overall genome improvement, given the advantages it has over the other conventionally used methods.

5.5. References

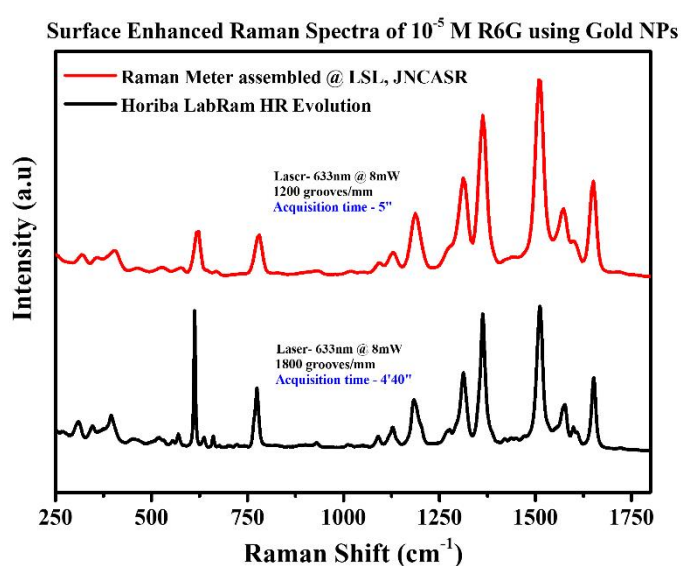
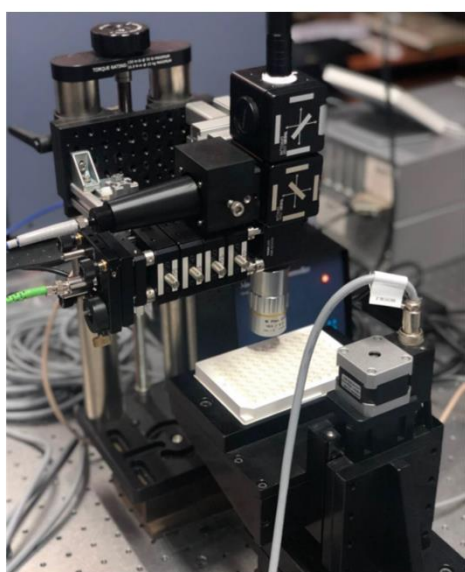
1. Czosnek, H. & Laterrot, H. A worldwide survey of Tomato Yellow Leaf Curl Viruses. *Archives of virology* 142, 1391 (1997).
2. Morin, S. et al. A GroEL Homologue from Endosymbiotic Bacteria of the Whitefly *Bemisia tabaci* Is Implicated in the Circulative Transmission of Tomato Yellow Leaf Curl Virus. *Virology* 256, 75 (1999).
3. Thongrit D, A.S., Sutabutra T. Tomato yellow leaf curl virus in Thailand-Plant Virus Diseases of Horticultural Crops in the Tropics and Subtropics. *FFTC Book Series* 33, 60 (1986).
4. Glick E., L.Y., Gafni Y. The viral etiology of tomato yellow leaf curl disease – a review. *Plant Protect. Sci.* 45, 81 (2009).
5. Ding, S.-W. & Voinnet, O. Antiviral Immunity Directed by Small RNAs. *Cell* 130, 413 (2007).
6. Bernstein, E., Caudy, A.A., Hammond, S.M. & Hannon, G.J. Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* 409, 363 (2001).
7. Verlaan, M.G. et al. Chromosomal rearrangements between tomato and *Solanum chilense* hamper mapping and breeding of the TYLCV resistance gene Ty-1. *The Plant Journal* 68, 1093 (2011).
8. Verlaan, M.G. et al. The Tomato Yellow Leaf Curl Virus Resistance Genes Ty-1 and Ty-3 Are Allelic and Code for DFDGD-Class RNA-Dependent RNA Polymerases. *PLOS Genetics* 9, e1003399 (2013).
9. Williamson, V.M. & Hussey, R.S. Nematode pathogenesis and resistance in plants. *The Plant cell* 8, 1735 (1996).
10. Sasser, J.N., Carter, C.C. Overview of the International Meloidogyne Project 1975-1984. *An advanced treatise on Meloidogyne - Biology and Control* 1, 19 (1985).
11. Zubeyir Devran, A.G., Lutfiye Mesci. Development of Molecular Markers for the Mi-1 Gene in Tomato Using the KASP Genotyping Assay. *Horticulture, Environment, and Biotechnology* 57, 156 (2016).
12. O'Neill R., S.R., Köhler W. Population Genetics: Aspects of Biodiversity. *Progress in Botany* 64, 115 (2003).
13. Tyagi, S. & Kramer, F.R. Molecular Beacons: Probes that Fluoresce upon Hybridization. *Nature Biotechnology* 14, 303 (1996).

14. Mhlanga, M.M., Malmberg, L. Using molecular beacons to detect single-nucleotide polymorphisms with real-time PCR. *Methods* 25, 463 (2001).
15. Santangelo, P.J., Nix, B., Tsourkas, A. & Bao, G. Dual FRET molecular beacons for mRNA detection in living cells. *Nucleic acids research* 32, e57 (2004).
16. Jothikumar, P., Hill, V. & Narayanan, J. Design of FRET-TaqMan probes for multiplex real-time PCR using an internal positive control. *BioTechniques* 46, 519 (2009).
17. Zhang, H. et al. Determination of Advantages and Limitations of qPCR Duplexing in a Single Fluorescent Channel. *ACS Omega* 6, 22292 (2021).
18. van Lierop, D., Faulds, K. & Graham, D. Separation Free DNA Detection Using Surface Enhanced Raman Scattering. *Analytical Chemistry* 83, 5817 (2011).
19. http://www.premierbiosoft.com/tech_notes/PCR_Primer_Design.html.
20. van Lierop, D., Larmour, I.A., Faulds, K. & Graham, D. SERS Primers and Their Mode of Action for Pathogen DNA Detection. *Analytical Chemistry* 85, 1408 (2013).
21. Cao YunWei, C., Jin, R. & Mirkin Chad, A. Nanoparticles with Raman Spectroscopic Fingerprints for DNA and RNA Detection. *Science* 297, 1536 (2002).
22. Faulds, K., Smith, W.E. & Graham, D. Evaluation of Surface-Enhanced Resonance Raman Scattering for Quantitative DNA Analysis. *Analytical Chemistry* 76, 412 (2004).
23. Hu, J. et al. Sub-attomolar HIV-1 DNA detection using surface-enhanced Raman spectroscopy. *Analyst* 135, 1084 (2010).
24. Wang, L., Roitberg, A., Meuse, C. & Gaigalas, A.K. Raman and FTIR spectroscopies of fluorescein in solutions. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* 57, 1781 (2001).
25. Sato, H., Kawasaki, M., Kasatani, K. & Katsumata, M.-a. Raman spectra of some indo-, thia- and seleno-carbocyanine dyes. *Journal of Raman Spectroscopy* 19, 129 (1988).
26. Basu, H.S. & Marton, L.J. The interaction of spermine and pentamines with DNA. *Biochemical Journal* 244, 243 (1987).

27. Lee, C., & 주재범. Selective Trace Analysis of Mercury (II) Ions in Aqueous Media Using SERS-Based Aptamer Sensor. *Bulletin of the Korean Chemical Society* 32, 2003 (2011).
28. Graham, D. et al. Selective Detection of Deoxyribonucleic Acid at Ultralow Concentrations by SERRS. *Analytical Chemistry* 69, 4703 (1997).
29. Jones, J.C. et al. Quantitative Assessment of Surface-Enhanced Resonance Raman Scattering for the Analysis of Dyes on Colloidal Silver. *Analytical Chemistry* 71, 596 (1999).
30. Lakhanpaul, T.J.a.S. Single nucleotide polymorphism (SNP)–Methods and applications in plant genetics: A review. *Indian Journal of Biotechnology* 5, 435 (2006).

Chapter 6. Construction of a miniaturised Raman spectrometer for applications in bio-detection

The instrumentation of Raman spectroscopy has been improved vastly over the past few decades. There have been many successful attempts at obtaining spectrometers that can provide better quality Raman spectra. Various spectral parameters like resolution, intensities and cutoff of the Raman bands have been improved, with better gratings, detectors and filters. Many modifications for performing *in situ* measurements have been implemented to study samples under varying external conditions like temperature, pressure and so on. Here, we have attempted to develop a miniature tabletop Raman spectrometer that can be used for point care diagnostic applications. It requires compactness, maximum signal collection and compatibility with well plates at low costs. We have successfully designed, assembled and tested the functioning of such a spectrometer. The spectrometer was assembled using components from various companies, and the motorized XYZ stage was automated to collect spectra from the various wells of a 96-well plate, thus reducing the overall cost of the Raman spectrometer.



6.1. Introduction

Raman spectroscopy is a scattering technique based on the Raman effect that measures the small fraction of inelastically scattered light when monochromatic radiation is incident on a sample. This technique probes the molecular vibrational and rotational frequencies. In Raman spectroscopy, a sample is illuminated with a monochromatic laser beam that interacts with the molecules to produce scattered light. The frequencies of this scattered light are mapped to construct a Raman spectrum. When the frequency of the scattered light is higher than the incident frequency, it is Stokes scattering, and when it is lower than the incident frequency, it is anti-Stokes scattering.

Depending on the incident laser wavelength, the frequency of the scattered light for a molecule varies. Hence, a Raman spectrum is always plotted as the Raman shift instead of frequency. The Raman shift is wavelength-independent. Since the Raman scattering of water is weak, it acts as a good solvent for dissolving samples.¹ Glass, and fused quartz also have weak Raman scattering, and hence they can be used for sample holders and optical components like mirrors and lenses.²

In the initial days of Raman spectroscopy, mercury arc lamps with a line of 435.8 nm were used as a light source. The discovery of the laser changed the course of Raman spectroscopy, as it provided a stable and intense beam of radiation. Today, there are many kinds of lasers that can be incorporated according to one's needs. The development of many optical components like the various filters, and gratings, drastically improve the quality of the spectral features, thus giving deeper insights into the vibrations of the molecules.

As seen previously in Chapter 5. , we have developed a strategy to detect indels and SNPs in the DNA extracted from the tomato plants. This strategy works on detecting the presence of either FAM or Cy3 dye-specific peaks in the resultant Raman spectrum. A compact Raman spectrometer with a high collection of scattered light serves the purpose of such fluorophore detections.

This motivated us to design a prototype for point of care diagnostics with the aim of a) developing the collection optics for maximum signal detection and b) developing detection techniques that are cheap to bring down the cost of the spectrometer without losing the high sensitivity of the system. Here, we demonstrate our plan of constructing a miniaturized table-top Raman spectrometer. The attempt is to develop a system that preferably considers the vicinity of a highly-specialized region of the Raman marker with respect to the noise rather than interpreting the whole spectrum. Hence the choice of the dispersing unit would be to give signal counts when it is aligned with the characteristic peak and give a background in the neighbouring spectral range. The signal to background ratio then helps quantify the concentration of the Raman marker.

6.2. Essential components of a Raman spectrometer

There are three essential things for constructing a Raman spectrometer: the incident light, the sample under study and the collection optics (Fig. 6.1) that gathers the scattered light and converts it into a readable spectrum on a software or as an image.

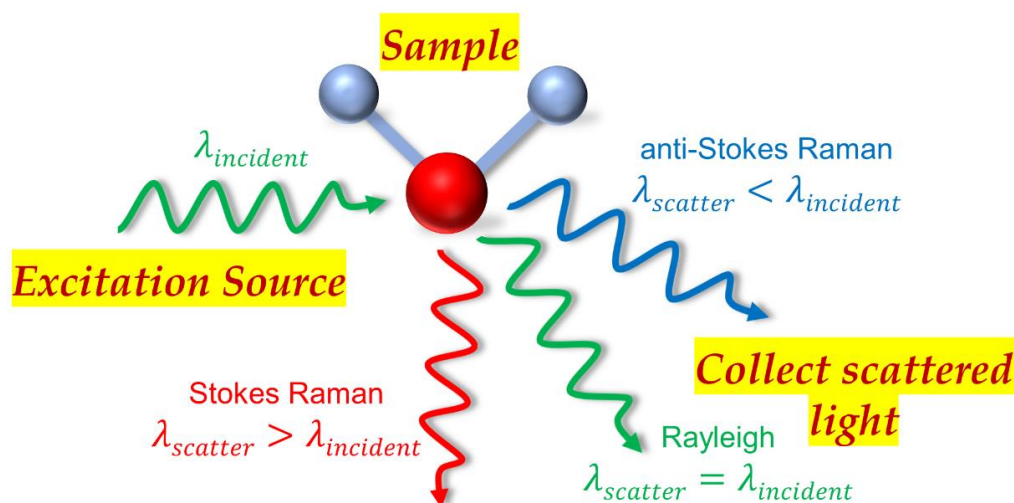


Fig. 6.1. Basic idea of Raman scattering – excitation source, sample and scattered light.

Raman spectrometers can be dispersive or non-dispersive. The dispersive Raman spectrometers use a diffraction grating or a prism to split the scattered light into its constituents. The non-dispersive Raman spectrometer uses an interferometer like a Michelson interferometer, like in the case of a Fourier transform Raman

spectrophotometer. The dispersive Raman spectrometer can detect lower wavenumbers, are sensitive, spatially well resolved, and operate at minimal costs. The non-dispersive Raman spectrometers have high spectral resolutions, fluorescence suppression and speed of analysis. Hence, dispersive spectrometers are most commonly used for standard detection and analysis applications.³

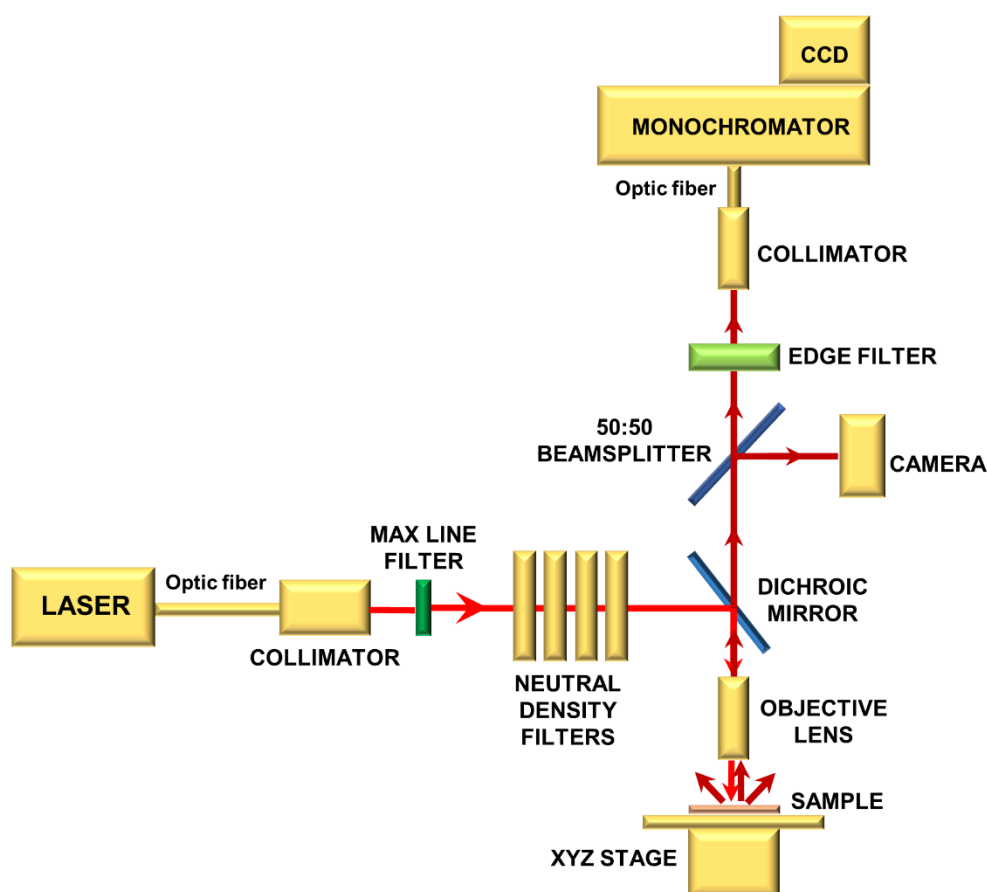


Fig. 6.2. Schematic representation of the components in a Raman spectrometer.

The multiple components that constitute a basic Raman spectrometer (Fig. 6.2) are discussed below:

a) Laser: Raman lines for a vibrational mode are narrow and highly specific. Thus to obtain a precise and proper width of the Raman line, the incident laser needs to have a stable wavelength position, bandwidth and spatial mode for accurate results. Continuous-wave (CW), single-longitudinal mode (SLM) and TEM₀₀ mode lasers are used for most applications.

When both the electric and magnetic field vectors are normal to the direction of propagation, they constitute the transverse electromagnetic (TEM) mode. In diode lasers, these are spatial modes perpendicular to the active layer. TEM₀₀ is the simplest transverse mode of a laser, and it has a Gaussian energy distribution across the beam. This allows for a tighter laser spot focused on the sample, resulting in good spatial resolution (Fig. 6.3a).

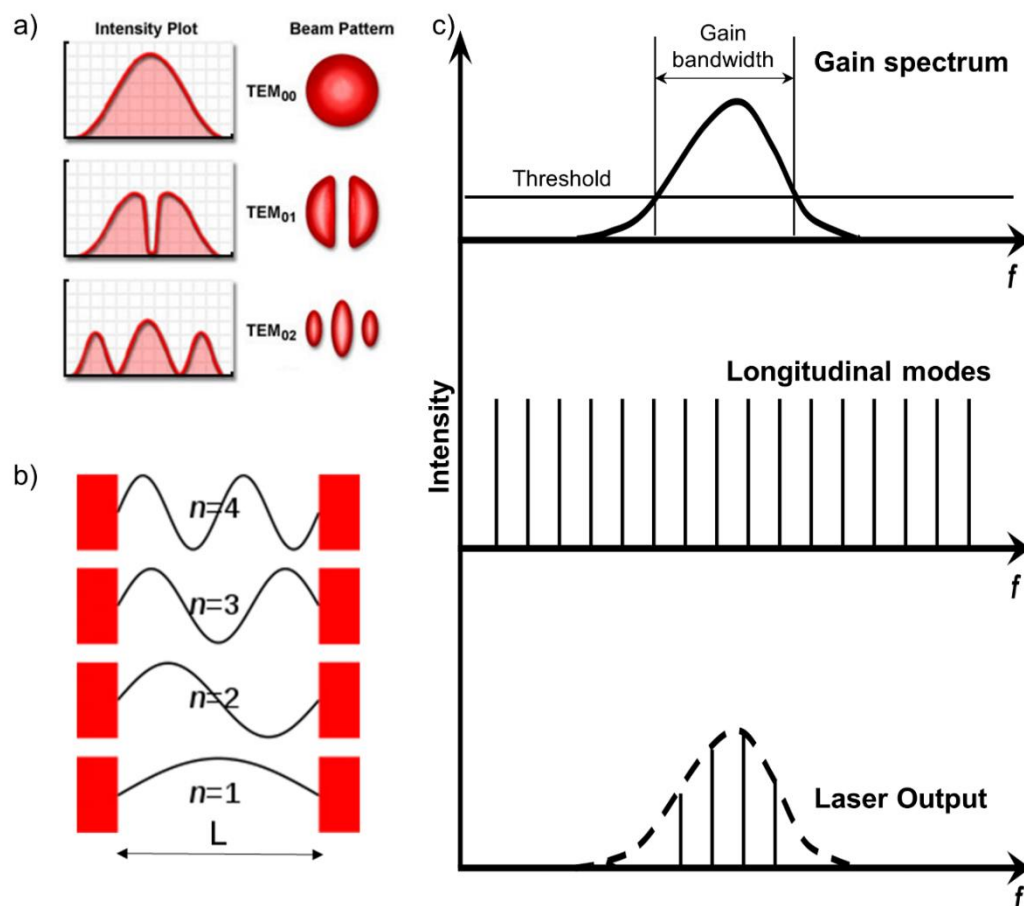


Fig. 6.3. a) A few TEM_{mm} modes, representing their energy distributions and the spatial profiles, b) various longitudinal mode in the cavity of the laser, c) the frequency content of a laser from its gain spectrum profile and its longitudinal modes.

When a standing wave is created in a cavity of a fixed length, it has a fundamental frequency of oscillation, represented by $n=1$ mode in Fig. 6.3b. There can be infinite overtones that exist, resulting in frequencies with discrete spacing between them and dependent on the geometry of the cavity. As seen in Fig. 6.3c, the laser has a specific gain spectrum with a threshold. In the middle is the representation of the longitudinal modes. If more than one mode is above the threshold, the laser is said to be a multi-longitudinal mode laser (as seen in Fig. 6.3c, bottom spectrum), and if only

one mode is above the threshold, it is a single longitudinal mode laser.⁴ For Raman spectroscopy, an SLM laser is essential as multiple excitation frequencies result in multiple scattered frequencies, thus causing a loss in spectral resolution due to these overlaps.

Today there are many kinds of lasers, including argon-ion laser (488 and 514.5 nm), helium-neon gas laser (632.8 nm), near IR diode lasers (785 and 830 nm), neodymium-yttrium aluminium garnet (Nd:YAG) (532 nm) and so on. Depending on the purpose of the study, a specific laser shall be chosen. The intensity of a Raman peak is inversely proportional to the fourth power of the laser's wavelength. Therefore, the intensity reduces as a higher wavelength laser is chosen. The higher the wavelength, the higher the penetration of the laser into the sample. The total penetration depends on the sample's absorptivity at a particular wavelength. If the sample under study is transparent, then the depth of focus and focal volume of the laser is given by the numerical aperture of the lens, wavelength of the laser and the sample's refractive index. Many molecules fluoresce when excited by wavelengths in the visible region. An optimal laser should be chosen to be at the tail of the fluorescence curve, hence reducing the background.

For the SERS experiments, the lasers not only have to excite the sample but also the surface plasmons of the metallic nanoparticles. Most silver and gold nanoparticles' surface plasmons are activated by visible light lasers such as 532 and 633 nm. As in SERS, the fluorescence gets quenched, and hence, the SPR activating lasers can be used. The miniature Raman spectrometer we are constructing is aimed at bioanalysis and bio-detection purposes. Most of such experiments are based on SERS, and hence the 532 and 632.8 nm lasers were chosen for the construction.

b) Max-line or laser-line filter: It is a precision optical filter with an extremely narrow passband centred on a standard laser wavelength (Fig. 6.4). They help screen out undesirable background radiation and ambient light. Laser-line filters were chosen according to the laser installed in the system.

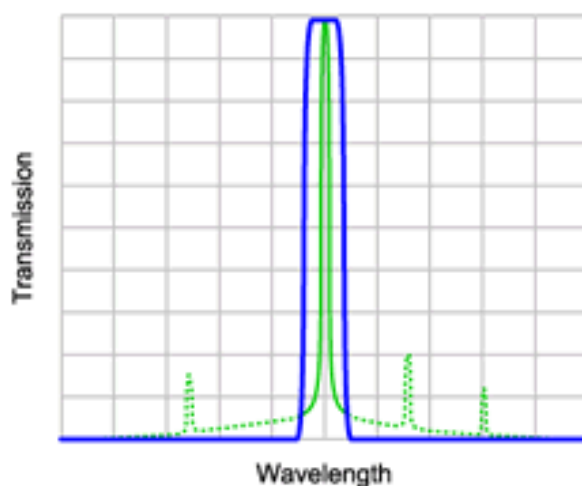


Fig. 6.4. Transmission curve of laser-line filter (blue), with the laser emission spectrum (green).

c) Neutral Density Filters: Neutral density (ND) filters can be absorptive or reflective. They are made from a family of glasses (Fig. 6.5a) that generally have a spectrally flat absorption coefficient (Fig. 6.5b). By varying the glass's type and thickness, a range of ND filters can be made with different absorption ranges. Generally, the optical density (OD) is measured for such glasses. It indicates the attenuation factor provided by an optical filter, i.e., how much it reduces the optical power of an incident beam. $OD = \log_{10}(1/T)$, where T is the transmission ranging between 0 and 1. Neutral density filters are available as individual filters of a specific OD or as a continuous wheel.

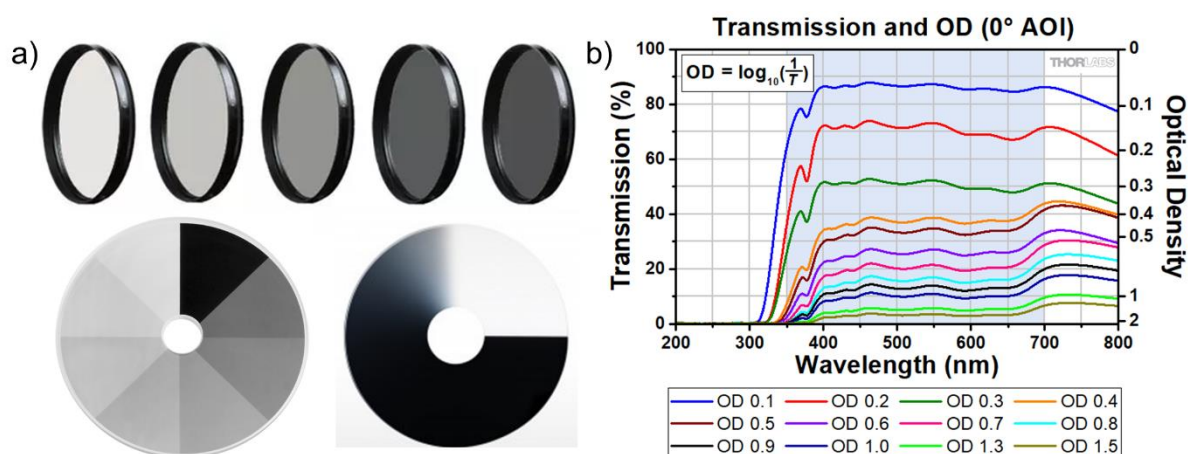


Fig. 6.5. a) Various kinds of ND filters – distinct, step-variable and continuously variable, and b) Transmission curves of ND filters of different ODs, blue region – wavelength range for anti-reflective coating.

d) Dichroic Mirror: These mirrors are coated such that they transmit one spectral region (passband) and reflect another spectral region (rejection band). This is

a short wave pass (SWP) if the passband is left of the transition slope and a long wave pass (LWP) if the passband is to the right of the slope (Fig. 6.6).⁵ Here the rejection band is reflected than absorbed. A steeper transition slope between passband and rejection band allows a sharper cut-on and cut-off. The coating can be customized to create pass- and rejection bands depending on the laser excitation wavelength.

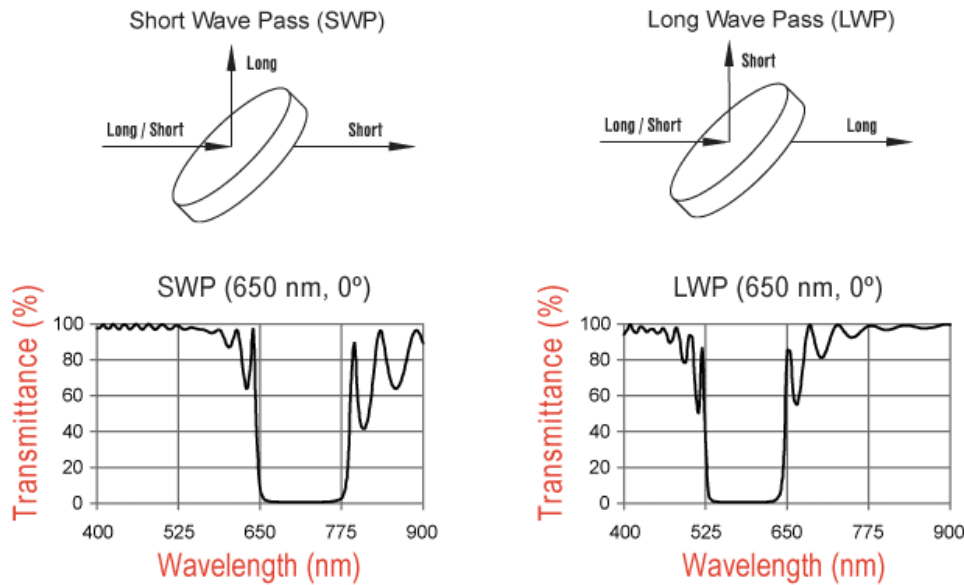


Fig. 6.6. The short-wave pass (SWP) and Long-wave pass (LWP) modes and their corresponding transmission curves.⁵

e) Objective: The objective allows the laser light to focus on the sample under study. The microscope objective’s magnification, clear and numerical aperture contribute to the efficient light collection.⁶

The clear aperture (CA) is the diameter of the rear aperture of an objective. It describes the limited light gathering of an optical system. The area is usually restricted to the edge or outer surface of the component. The numerical aperture (NA) depends on the medium’s refractive index through which the light passes, n , as $NA = n \sin \theta$ where θ is the angle between the ray along the optical axis at the centre of the lens and the ray from the perimeter of the lens to the sample. Higher the NA , the greater the solid angle of light collection.⁶

The spot size that a laser can be focussed down to in Raman spectroscopy depends on the laser wavelength and the properties of the objective lens. The central

part of the airy disc constituting 84 % of the intensity is the laser spot size given by $1.22\lambda/NA$. With shorter laser wavelengths and higher NA values, smaller spot sizes can be achieved, thus improving the spatial resolution as seen in Fig. 6.7.⁷

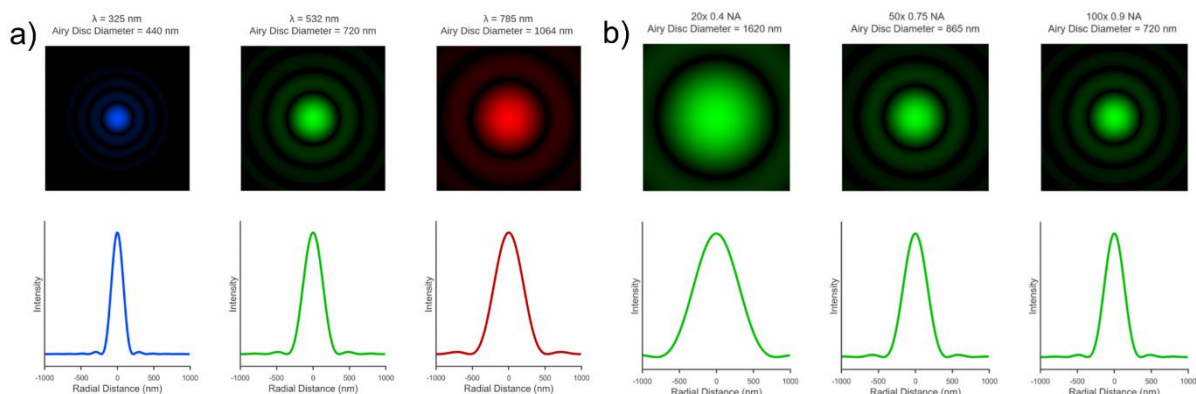


Fig. 6.7. Dependence of the laser spot size of the sample on a) laser wavelength and b) the magnification and numerical aperture of the microscope objective.⁷

f) Edge Filter: The edge filter splits the incoming light into two components. Wavelengths lesser than a cut-off are reflected or absorbed, and those higher than the cut-off are transmitted (Fig. 6.8a). This allows us to remove the anti-Stokes and the Rayleigh scattered light. To retain the anti-Stokes scattering, a notch filter (a bandpass filter) is chosen (Fig. 6.8b) instead of an edge filter.

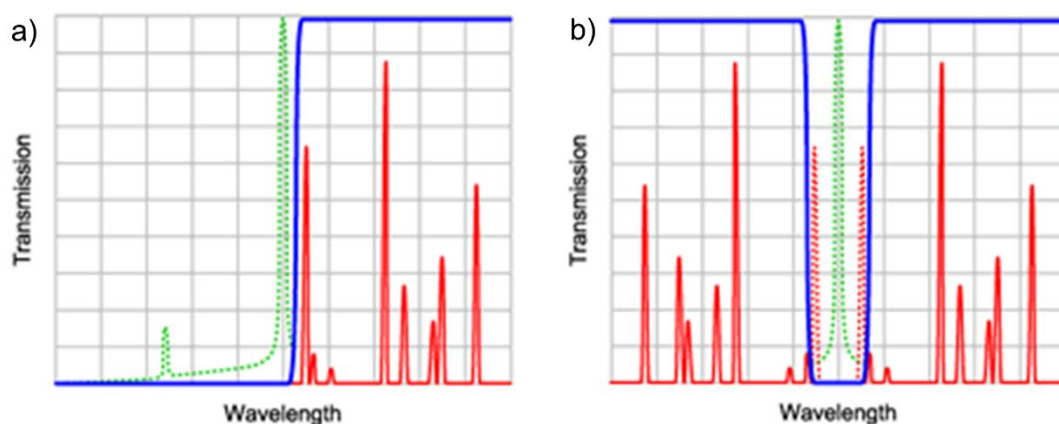


Fig. 6.8. The transmission curve (blue) of the a) Edge filter and b) Notch filter. Laser line and scattered Raman light are represented in green and red, respectively.

For most spectrometers that involve the studies of specific Raman spectral features, collecting the Stokes scattered light is sufficient, and hence the edge filter can be used. The edge filter has an advantage, where its transmission curve can be tuned to be ultra-step, thus obtaining peaks at very low Raman shifts.

g) Beamsplitter and camera: A 50:50 beamsplitter diverts 50 % of the light onto a camera to capture the white light image of the sample, along with locating the position of the laser spot on the sample. This setup is beneficial but optional.

h) Spectrograph: A monochromator separates the incoming radiation into its constituent monochromatic wavelengths. A dispersive unit, either a prism or a diffraction grating, is used to create these monochromatic lights. Change in refractive index as a function of wavelength helps a prism separate the incoming radiation into its components. A diffracting grating has serrated grating lines separated by a fixed distance. These lines cause the light to diffract and split into a spectrum. When the split spectrum is incident on an electronic detector to obtain a spectrum, the setup is called a spectrograph.

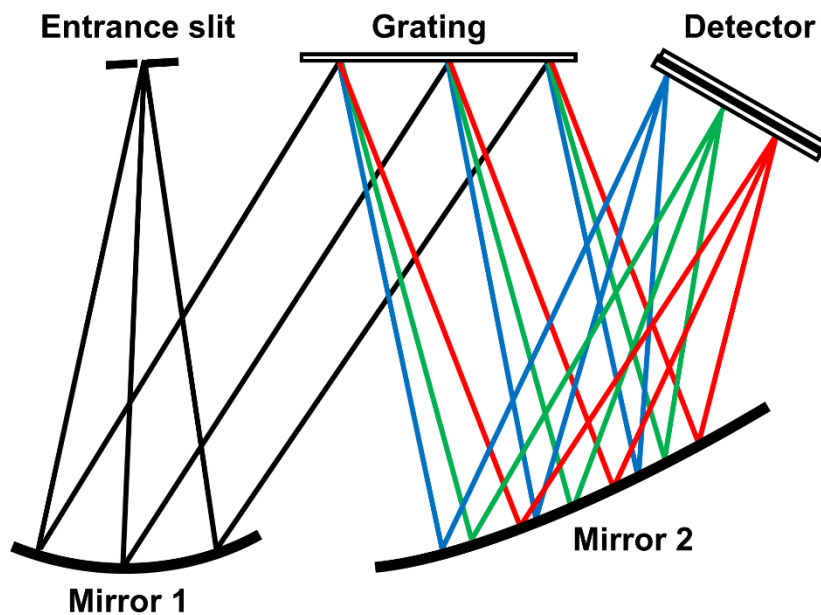


Fig. 6.9. Schematic representation of the Czerny-Turner configuration of the spectrometer.

There are many kinds of spectrometers that have been developed over the years. Some of them include the Rowland circle construction using a concave grating, the Czerny-Turner configuration using two mirrors, the Fastie-Ebert construction using one large spherical mirror and the Sargent-Rozey version using two tilted mirrors. The most commonly used configuration is the one given by the Czerny-Turner (Fig. 6.9).⁸
⁹ It consists of two concave mirrors and one planar diffraction grating. Mirror 1 collimates the light source onto the grating, and mirror 2 focuses the dispersed light

from the diffraction grating onto the CCD detector. The larger the focal length of the monochromator, the better resolution of the spectral peaks.

i) Various geometries of Raman spectroscopy:

There are two main setups through which the Raman spectra can be obtained, a) 180° backscattering geometry (Fig. 6.10a) and b) 0° transmission geometry (Fig. 6.10b). There is a third geometry that has been explored in the recent past, called the spatially offset Raman spectroscopy mode (Fig. 6.10c).¹⁰

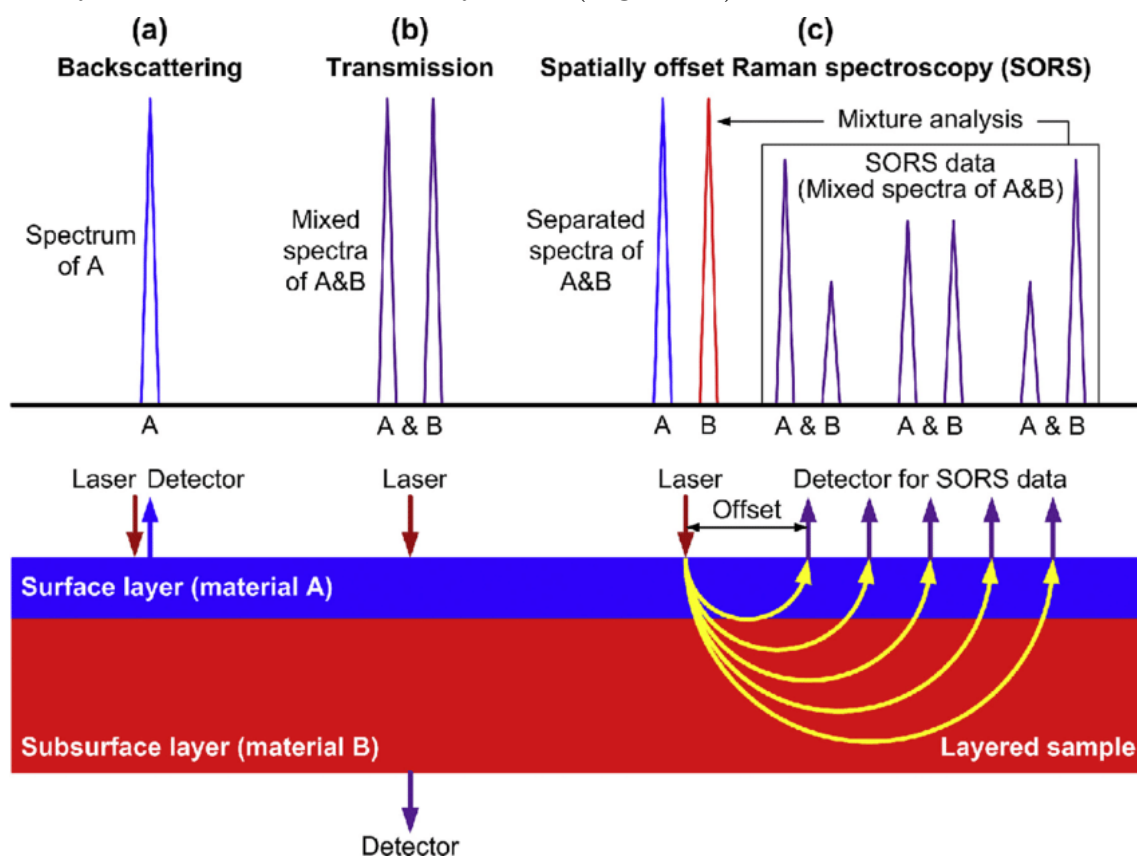


Fig. 6.10. Schematic representation of the Raman measurement modes, a) Backscattering, b) Transmission, c) spatially offset Raman spectroscopy.^{†††}

In the earlier days of the development of Raman spectrometers, transverse geometry (90° angle between the illumination and collection optics) was used, and is still being used for liquid and gaseous samples in the present day. But over the days, the backscattering geometry was found to be more advantageous, as it allows multiple

^{†††} Reprinted with permission from Subsurface inspection of food safety and quality using line-scan spatially offset Raman spectroscopy technique, 75, Jianwei Qin et al, 246-254, Copyright (2017) from Elsevier.

spectroscopy studies simultaneously, makes it easy to work with single crystals and greater energy distribution and hence lower sample heating.^{11, 12} Today, the 180° backscattering geometry is one of the most commonly used setups for a Raman spectrometer. In this geometry, the laser and the detector are on the same side of the sample.

The transmission Raman spectroscopy is a variant of Raman spectroscopy in which the sample is illuminated from the top, and the scattered light is collected in the same direction. This method is advantageous in probing the bulk content of diffusely scattered samples. It has been used to probe the bulk content of powders and tissues and has been able to reject Raman and fluorescence components originating from the surface of the sample.¹³

The spatially offset Raman spectroscopy (SORS) is a variant of Raman spectroscopy, which probes the scattered light away from the point of incidence of the laser light. This allows for highly accurate chemical analysis of objects below tissues, coatings, bottles, and so on.¹⁴

6.3. Construction of the miniature Raman spectrometer

From the above knowledge about the various components required to build a Raman spectrometer, we constructed a compact Raman spectrometer (Fig. 6.11) with the following specifications for the various components.

a) Laser – A 632.8 nm SLM diode laser with TEM₀₀ mode was chosen for the SERS based bio-application studies. The diode lasers have very low power consumption, are compact, have high output power stability and are relatively lower in cost. The source power is in the range of 20-50 mW, with a deviation of ± 0.5 nm in the central wavelength. The typical spectral linewidth is < 10 MHz.

b) Collimator – A collimator of focal length 24.98 mm is chosen to collimate the incoming laser radiation from the source through an optic fibre. It also acts as a beam expander, allowing the laser beam to expand, hence more tightly focused onto

the sample. The $1/e^2$ output of the beam diameter is 4.7 mm. This collimator is mounted onto a kinematic mount which allows for $\pm 5^\circ$ tip or tilt and a $\pm 3^\circ$ linear translation along the optic axis. This allows for fine-tuning the alignment of the incident laser onto the sample.

c) Max-line or laser line filter – A laser line filter with a central wavelength of 632.8 nm and a nominal bandwidth of 2.4 nm was used.

d) Neutral Density Filters – Four Neutral Density (ND) filters with optical densities (OD) of 0.1, 0.3, 0.5 and 1.0 in the visible range were installed. Depending on the power requirement, either a single or a combination of these can be chosen.

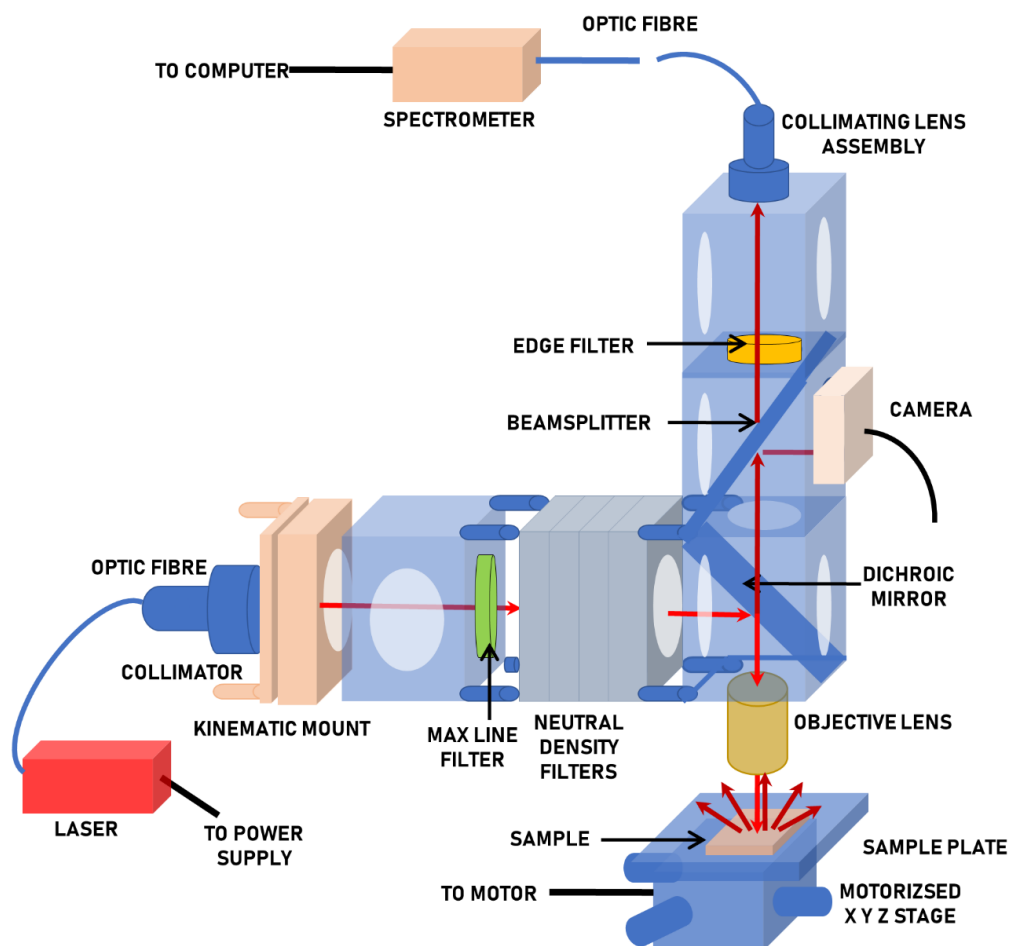


Fig. 6.11. Schematic representation of the assembled spectrometer.

e) Dichroic Mirror – A dichroic mirror for 632.8 nm central laser wavelength was chosen. It has an edge steepness of 3.2 nm, equivalent to 78.6 cm^{-1} . It has a

reflection, $R > 94\%$ for wavelengths below 632.8 nm, and a transmission, $T > 93\%$ for wavelengths in the range of 641 – 980.8 nm.

f) Objective – A long working distance (33.5 mm) infinity-corrected 10X objective with a large clear aperture was chosen to focus the light on the sample. The scattered light is collected back through this objective through the back-scattering mode.

g) Beam Splitter and Camera– A 50:50 Beam Splitter is chosen to reflect 50% of the scattered light onto a 5MP CMOS type camera when white light imaging of the sample and the laser spot on the sample is required. This beam splitter can be removed when Raman spectra are collected to allow all the light transmitted from the dichroic to pass further onto the edge filter.

h) Edge Filter – A 632.8 nm edge filter at an angle of incidence $0^\circ (\pm 2^\circ)$ is used to cut off the anti-Stokes and Rayleigh scattered light. It is a long pass edge filter with an edge wavelength of 634.5 nm. It blocks the Rayleigh light of 632.8 nm with an OD > 6 . Its transition width of 3.2 nm, equivalent to 79 cm^{-1} , allows for easy measurements of Raman shifts above 100 cm^{-1} .

i) Spectrograph – Through a custom-built collimating assembly, the light that passes through the edge filter is incident onto the pinhole of a commercially bought monochromator. The monochromator is integrated with a back-thinned, TE cooled, 1044x64 pixel CCD detector to record the Raman spectra. The spectrograph is compact and houses a 1200 grooves/mm holographic diffraction grating. It has a good signal to noise ratio of 1000:1. It can record a Raman spectrum between $150\text{-}3150 \text{ cm}^{-1}$ with a resolution of $7\text{-}10 \text{ cm}^{-1}$.

j) Motorized XYZ Stage – A custom-designed motorized XYZ stage with x- and y- travel of 100 mm and z- travel of 25 mm is used to manipulate the position and adjust the sample position to the focal point of the objective. The sample stage is custom made to accommodate a 96-well plate on it.

k) Cage Assembly – Various individual cage components, including cage cubes, connecting rods, cubes for mounting dichroic mirror and beam splitter, Neutral density filter holders and so on, have been used to put together all the components of the miniature Raman spectrometer. The compact arrangement of all these components assures minimal entry of stray light into the light's path inside the spectrometer and additionally reduces losses.

The list of components for assembling the spectrometer are given in Table. 6.1. All the components mentioned above were purchased from different companies and were assembled in-house. Choosing various components from different companies allowed for a greater choice in options, and a lower cost of the assembled spectrometer.

Table. 6.1. List of components for spectrometer assembly.

Component	Company (Model No.) Setup with 633 nm laser	Company (Model No.) Setup with 532 nm laser
Laser	PD-LD (SLM-632.8-H-PMF)	Integrated optics (0532L-41B-NI-NT-NF)
Collimator	Thorlabs (TC25APC-633)	Thorlabs (GBE05-A)
Laser Line Max Filter	Semrock (LL01-633-12.5)	Semrock (LL01-532-12.5)
Neutral Density Filters	Thorlabs (NExB-A series)	Thorlabs (NExB-A series)
Dichroic Mirror	Semrock (LPD02-633RU-25x36x1.1)	Semrock (LPD02-532RU-25x36x1.1)
Objective lens	Mitutoyo (No. #46-144)	Mitutoyo (No. #46-144)
Beamsplitter	Edmund Optics (32269)	-
Camera	MOTICAM (Moticam 5)	-
Edge Filter	Semrock (LP02-633RE-25)	Semrock (LP03-532RE-25)
Spectrograph	Ocean Optics (QEPRO-RAMAN-SPL)	Avantes (AvaSpec-ULS4096CL-EVO)
Motorized XYZ Stage	Holmarc	Holmarc
Cage Assembly	Thorlabs	Thorlabs

Similarly, a 532 nm excitation miniature Raman spectrometer was built with a more compact design idea (Fig. 6.12). This system acted as an upgraded version of the previous one regarding compactness and ease of accessibility.

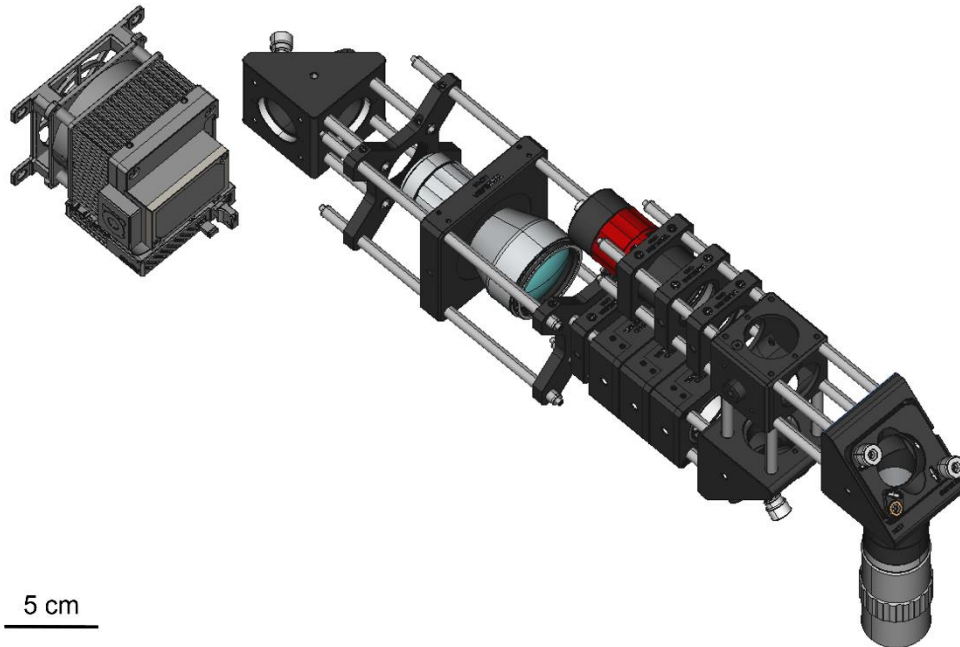


Fig. 6.12. Schematic representation of the optical component assembly for the 532 nm laser excitation setup.

Comparing the two kinds of setups, the model built with the 633 nm excitation laser had fixed cage cubes and a kinematic mount that allowed for fine-tuning the alignment of the laser spot. This allowed for easy assembly of the system. But also, small micron to mm size differences in the alignment of the cage cube components would easily misalign the system, which had to be then realigned frequently. While in the later model, there were 3 movable mirrors with an angle of incidence at 45° , allowing for fine-tuning the alignment at multiple levels, thus holding the alignment stable for long durations. Though the alignment times can be compared, most of the components are already in fixed positions for both these systems. This allows for easy laser alignment that can be done in a few minutes at the maximum.

The optical elements were installed in the respective optomechanical components. The installed opto-mechanics was assembled to create the optical path of the light for illuminating the sample and collection of the scattered light through 180° backscattering geometry (Fig. 6.13). Using the kinematic mount in the 633 nm setup

and the mirror mounts in the 532 nm setup, the laser light was aligned such that the incident laser spot overlapped the light in the collection path, indicating complete laser alignment. The assembled 633 nm laser spectrometer was mounted onto a vertical translation stage for coarse movement of the complete spectrometer. The sample XYZ stage allowed fine movements of the sample to align under the microscope objective. The assembled 532 nm laser spectrometer was placed on a vertical platform at an optimal height to accommodate the XYZ sample stage.

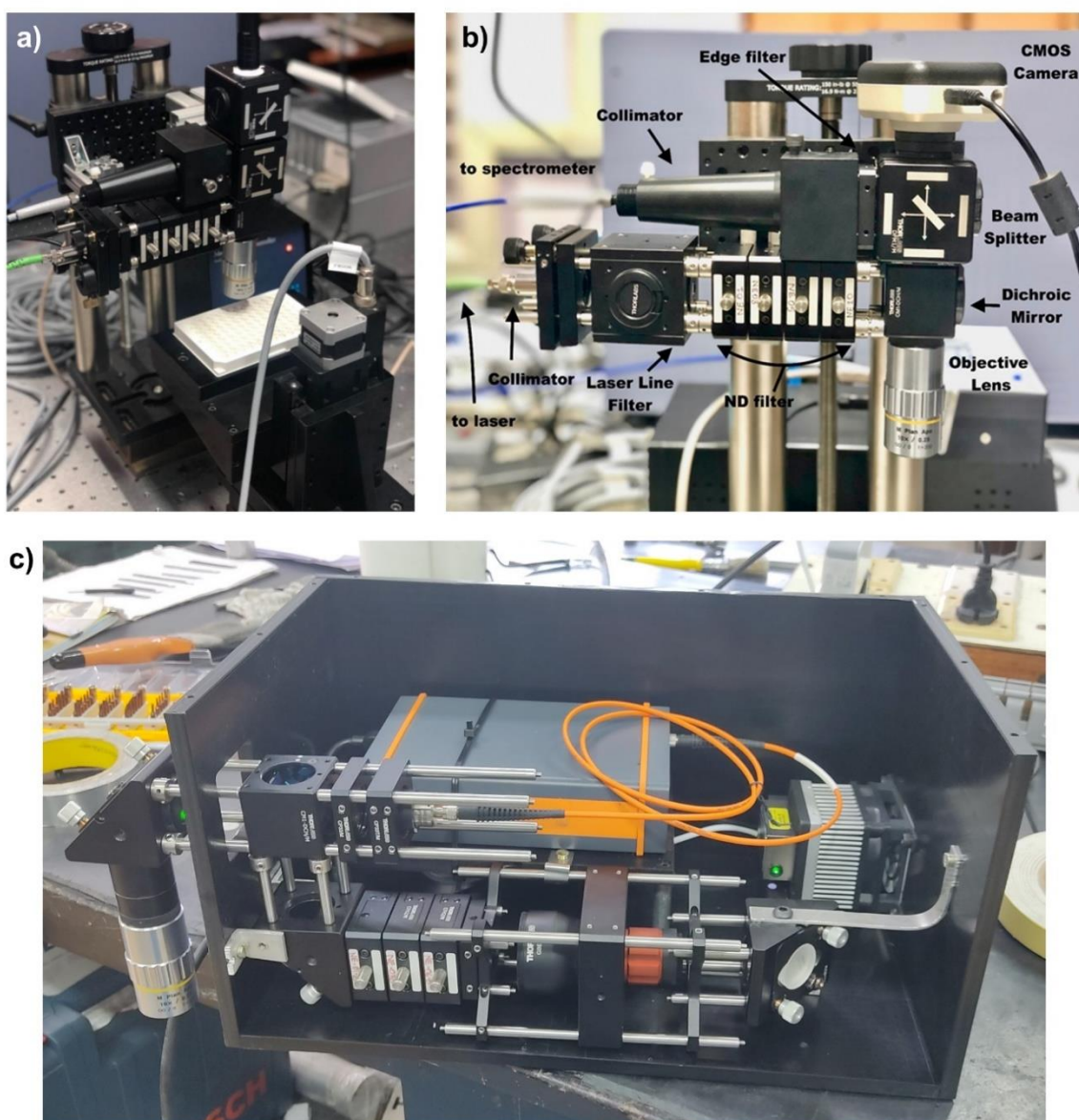


Fig. 6.13. a) Setup with 633 nm laser and a motorized XYZ stage to accommodate a 96-well plate, b) representation of the various components in the setup shown in a), and c) compact arrangement of all components in the 532 nm laser setup.

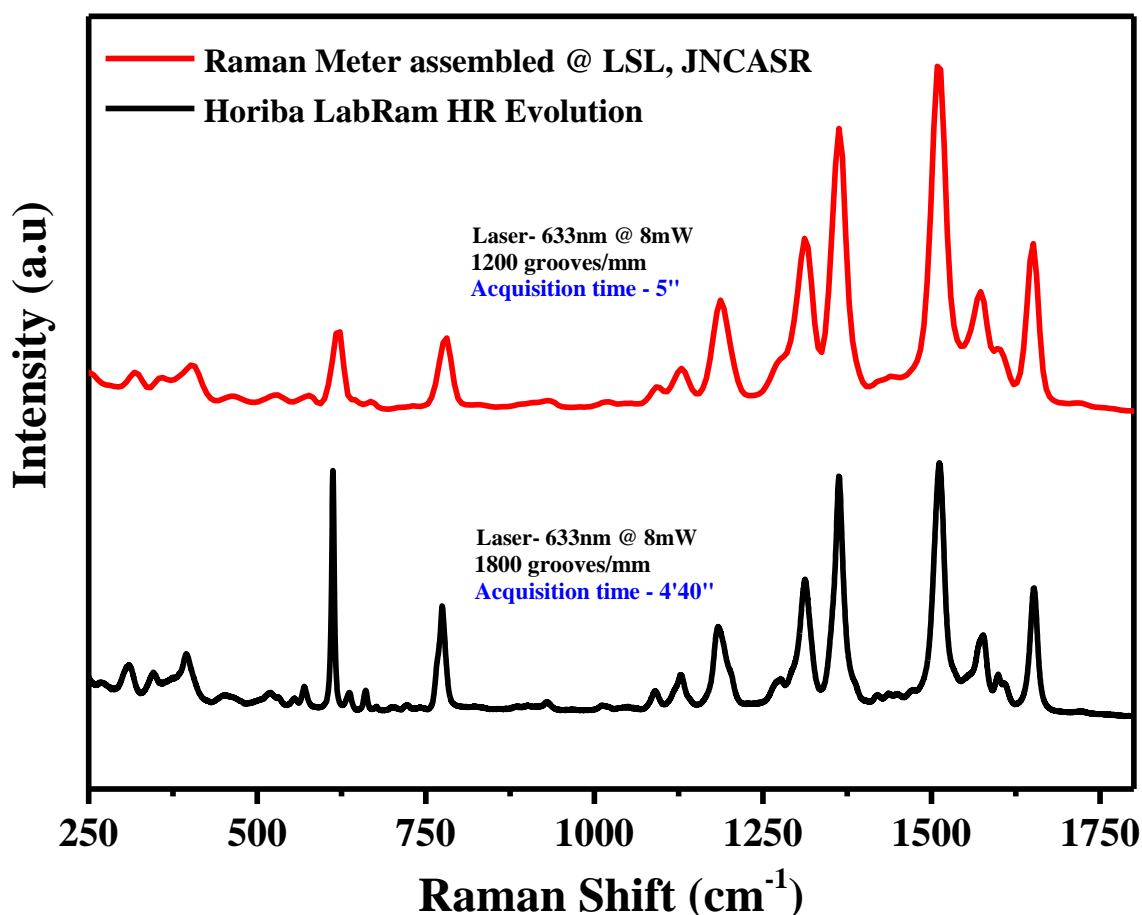


Fig. 6.14. Comparison of the SERS spectra of 10^{-5} M R6G recorded on the 633 nm laser spectrometer built in-house vs. the commercially available LabRam HR Evolution.

The SERS spectra for 10^{-5} M R6G using silver nanoparticles were recorded on the built spectrometer. As seen in Fig. 6.14, the SERS spectra were collected on the in-house assembled Raman spectrometer and the commercially available LabRam HR Evolution. The in-house built spectrometer produces a spectrum of similar intensity in just 5 s of time, while it took about 280 s on the LabRam HR Evolution. The spectrum on the LabRam has an excellent resolution, producing peaks of narrow FWHM (overall spectrum) and well-separated peaks (For example, peaks around $625\text{-}675\text{ cm}^{-1}$). This is attributed to the low $7\text{-}10\text{ cm}^{-1}$ resolution of the monochromator of the inhouse built spectrometer. For detection-based applications, when unique spectral regions are chosen, this low resolution doesn't interfere, and thus this spectrometer serves the purpose. The collimating assembly on the collection optics end can always be integrated to a higher focal length spectrograph to obtain spectra with higher resolution for research-based studies.

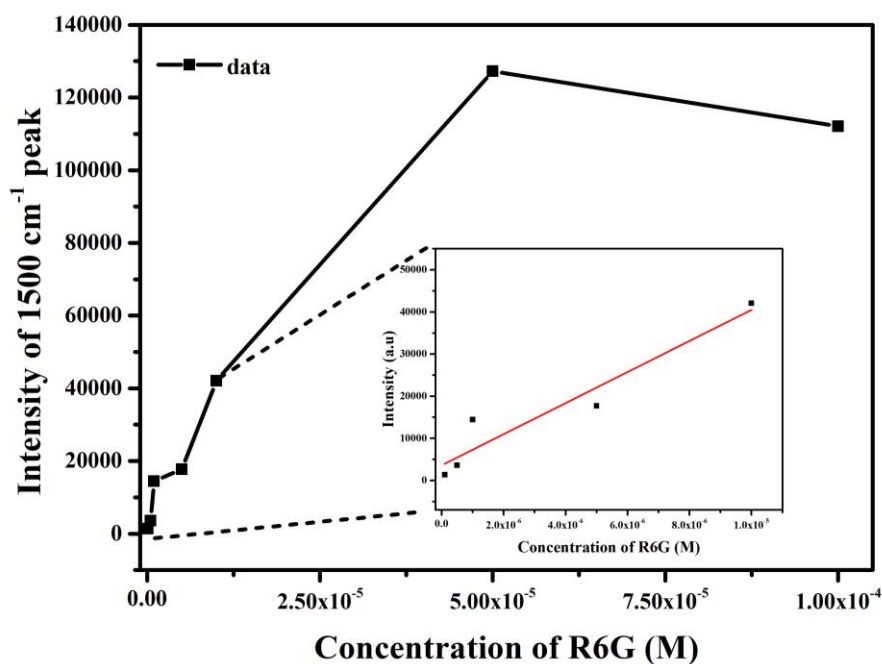


Fig. 6.15. Intensity of the 1500 cm^{-1} peak plotted as a function of the concentration of the R6G in the SERS spectra of R6G using AgNPs.

The SERS spectra of various concentrations of R6G were obtained using AgNPs synthesized by the Lee-Meisel method (Fig. 6.15). A least concentration of 10^{-8} M R6G was detected using the in-house built spectrometer. Such results give the confidence to detect low concentrations of dye tags in various applications.

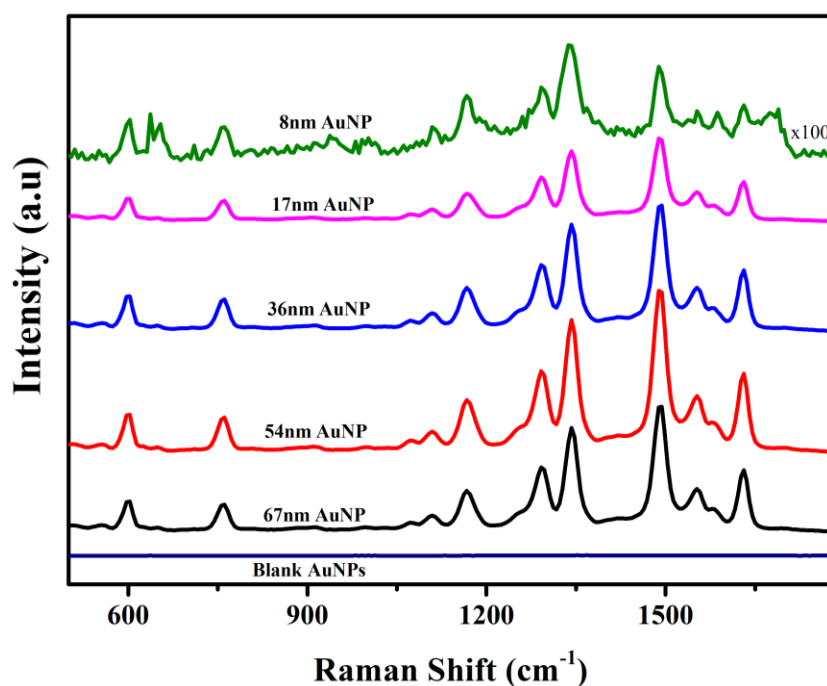


Fig. 6.16. Size-dependent SERS spectra of 10^{-5} M R6G using AuNPs.

As described in section 2.7, varying sizes of AuNPs were synthesized, and size-dependent SERS spectra of 10^{-5} M R6G were obtained. From Fig. 6.16, it is evident that the larger the AuNP size, the higher the intensity of the SERS spectrum. The optimal size range of AuNPs for best results is 50-60 nm, and for larger sizes, the SERS intensity reduces due to bulk effects in the nanoparticles.

With these sets of experiments, the functioning of the spectrometer and its ability to obtain SERS spectra for low concentrations of dyes was established. To make the spectrometer user-friendly for diagnostic labs, we developed a program to allow the XYZ motorized stage to map a 96-well plate and obtain spectra at each well. As seen in Fig. 6.17, a LabVIEW code was written to move specific distances at fixed intervals and wait at each position for a fixed period.

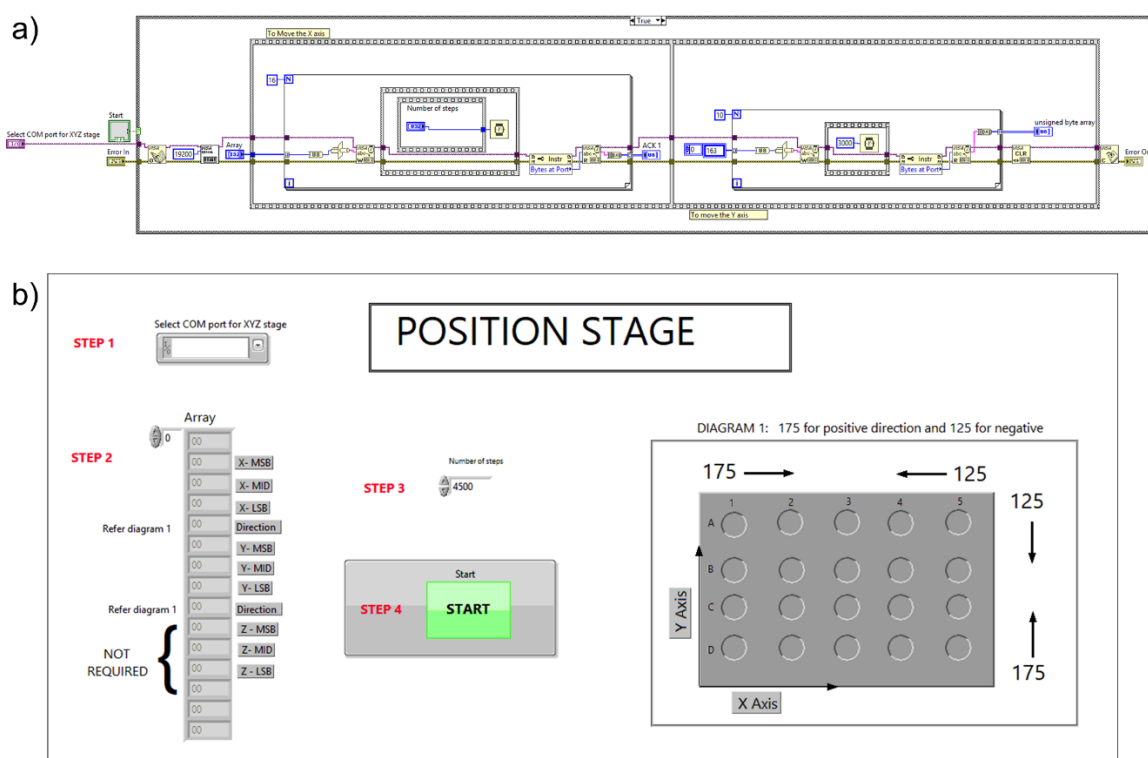


Fig. 6.17. a) Back panel of the LabVIEW program to position a specific well from the 96-well plate under the microscope objective, b) Front panel of the stage-positioner.

This program was then integrated with the Oceanview software corresponding to the spectrograph (Fig. 6.18a). It communicates to the spectrograph to collect spectra at optimal time points. The spectra were then displayed on the front panel designed on LabVIEW (Fig. 6.18b).

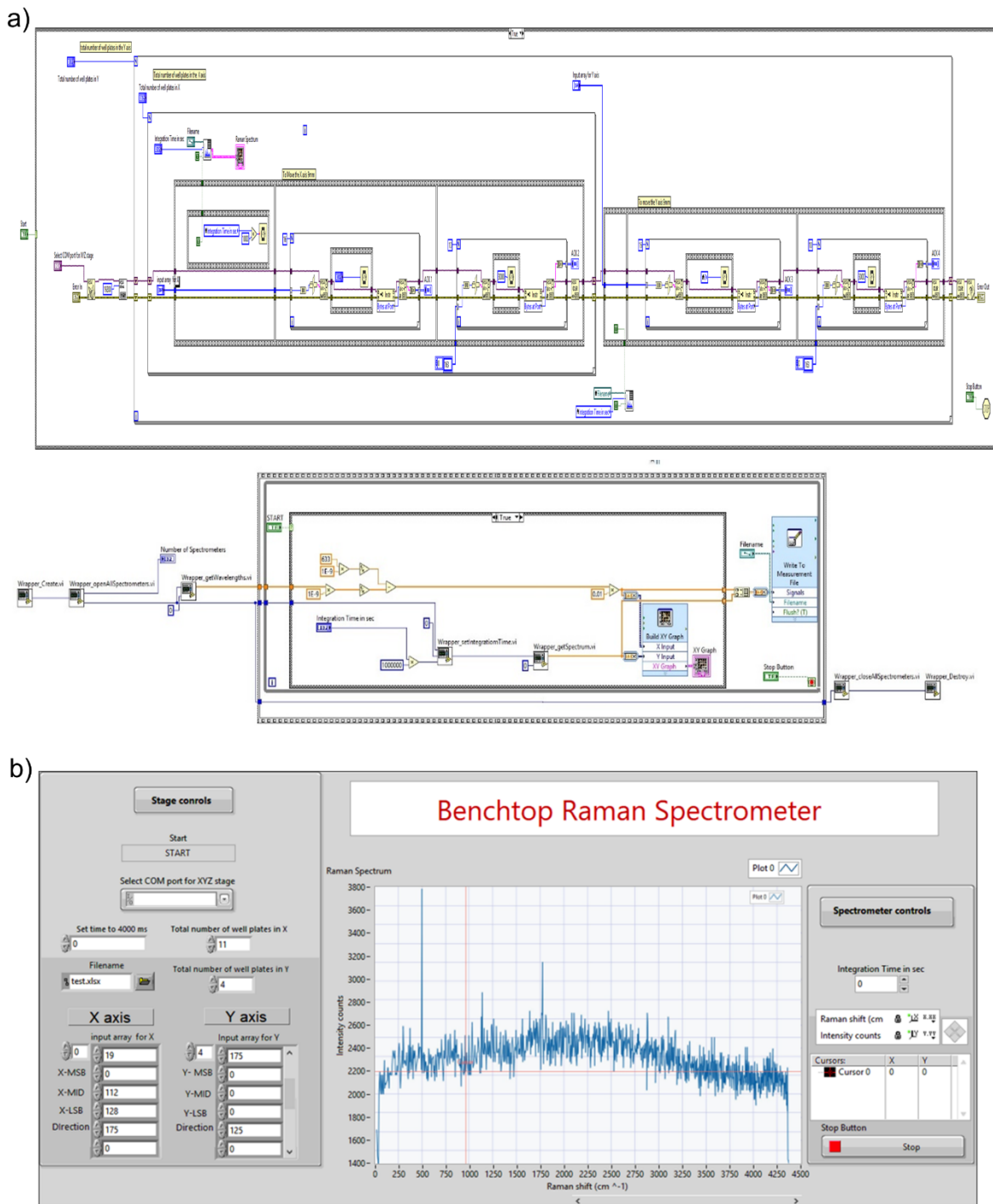


Fig. 6.18. a) Back panel views of the LabVIEW program to integrate the XYZ positioner to the Oceanview software of the spectrograph, and b) Front panel of the LabVIEW code where the stage can be positioned, and the spectrum can be collected.

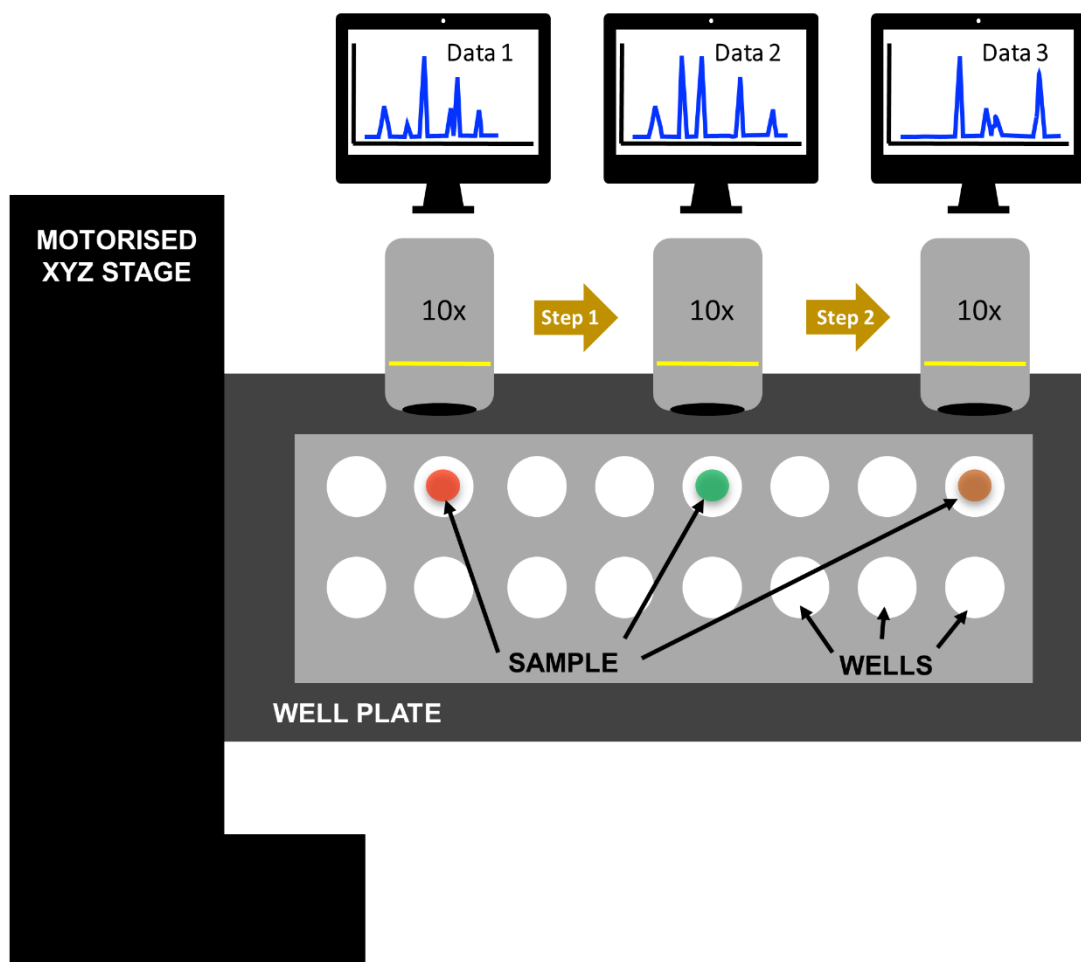


Fig. 6.19. Schematic representation of the automated movement of the XYZ stage to collect spectra in different wells of the 96-well plate.

In Fig. 6.19, a schematic representation of the movement of the XYZ stage to various well positions and the simultaneous collection of the SERS spectra is represented. The functioning of the motorised XYZ stage for the 96-well plate was tested by putting various Raman active liquid samples in the wells and collecting their spectra. We successfully collected the spectra of samples like toluene, cyclohexane, a mixture of AgNPs and 10^{-7} M R6G placed in different plate wells (Fig. 6.20). Hence, a successful assembly and programming of a motorized XYZ stage to an in-house Raman spectrometer was attained and tested.

The size (l x b x h), weight and cost of the 532 nm and 633 nm spectrometers are 25 cm x 35 cm x 18 cm, 50 cm x 30 cm x 45 cm; ~ 6 kg total, ~ 5 kg for the optics

and the monochromator, and an additional 8kg for the motorized xyz stage and its controller; and ~ 12-13 lakh INR, ~ 22 lakh INR, respectively.

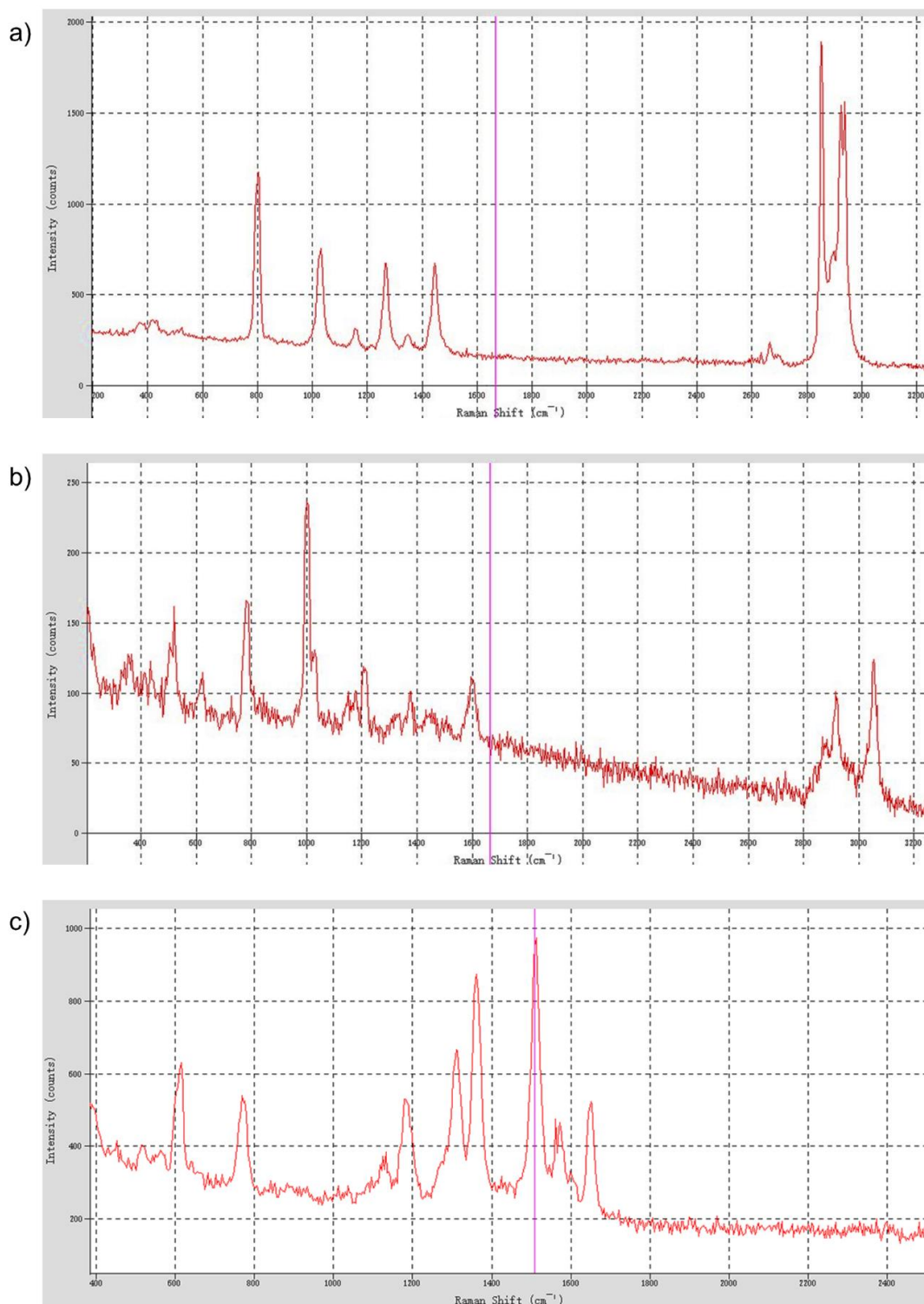


Fig. 6.20. Raman spectra were collected using the 633 nm laser Raman spectrometer with the program written for automated spectral collection at different wells, a) well 1 – cyclohexane, b) well 2 – toluene, and c) well 3 – AgNPs with 10^{-7} M R6G.

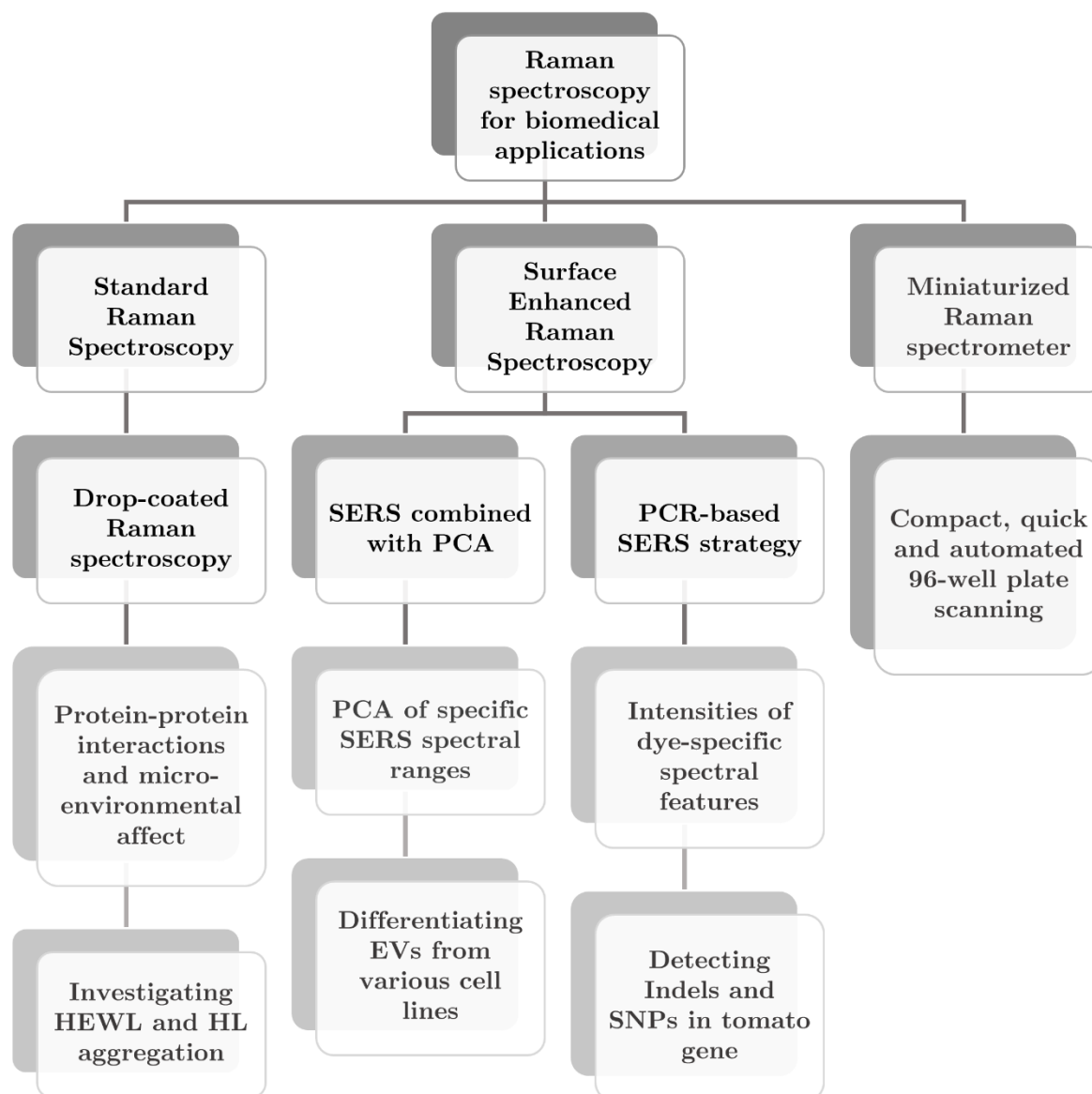
6.4. Conclusion

In this chapter, two miniature table-top Raman spectrometers with 532 and 633 nm laser excitation and different optomechanical components were assembled and evaluated. The spectrometers are compact and can be assembled and aligned easily on-site. Compared to a commercial research-grade spectrometer as a LabRam HR Evolution, these in-house built spectrometers were cheaper, compact in dimensions, and had quick integration times. As of today, due to the requirement of the product, it incorporates a nominal resolution, single wavelength, and single diffraction grating. According to the study, these parameters can be changed any day by integrating a large focal length monochromator, multiple lasers, and their corresponding optics. The advantage of this kind of setup includes a mix and match of optomechanical elements, allows for the system modification any day, and lets you build in other experimental features, for example, including the fluorescence measurement setup. Basic concentration-dependent SERS of rhodamine 6G was conducted as a study on the spectrometer. The custom-built XYZ stage was programmed to accommodate a 96-well plate and collect spectra at each well at a very nominal cost compared to the existing motorised stages available in the market. Overall, a tabletop Raman spectrometer for diverse applications was assembled and its functionality tested.

6.5. References

1. Walrafen, G.E. & Blatz, L.A. Weak Raman bands from water. *The Journal of Chemical Physics* 59, 2646 (1973).
2. Tobin, M.C. & Baak, T. Raman Spectra of Some Low-Expansion Glasses. *Journal of the Optical Society of America* 58, 1459 (1968).
3. Marleen de Veij, D.P.V., Prof. Dr. Luc Moens. The rise of Raman. <https://www.europeanpharmaceuticalreview.com/article/3018/the-rise-of-raman/> (2005).
4. <https://www.rpmclasers.com/what-is-single-longitudinal-mode/>.
5. <http://rmico.com/visible-near-ir-dichroic-filters>.
6. Tuschel, D. The Effect of Microscope Objectives on the Raman Spectra of Crystals. *Spectroscopy* 32, 14 (2017).
7. <https://www.edinst.com/blog/laser-spot-size-in-a-microscope/>.
8. Czerny, M., Turner, A.F. *Z.Physik* 61, 792 (1930).
9. Mohammadi, H. & Eslami, E. Investigation of spectral resolution in a Czerny Turner spectrograph. *Instruments and Experimental Techniques* 53, 549 (2010).
10. Qin, J. et al. Subsurface inspection of food safety and quality using line-scan spatially offset Raman spectroscopy technique. *Food Control* 75, 246 (2017).
11. Ferraro, J.R., Nakamoto, K. & Brown, C.W. in *Introductory Raman Spectroscopy (Second Edition)* (eds. Ferraro, J.R., Nakamoto, K. & Brown, C.W.) 95 (Academic Press, San Diego, 2003).
12. Shriver, D.F. & Dunn, J.B.R. The Backscattering Geometry for Raman Spectroscopy of Colored Materials. *Applied Spectroscopy* 28, 319 (1974).
13. Sparén, A. & Svensson, O. in *Encyclopedia of Spectroscopy and Spectrometry (Third Edition)* (eds. Lindon, J.C., Tranter, G.E. & Koppenaal, D.W.) 469 (Academic Press, Oxford, 2017).
14. Parker, A.W. in *Encyclopedia of Spectroscopy and Spectrometry (Third Edition)* (eds. Lindon, J.C., Tranter, G.E. & Koppenaal, D.W.) 143 (Academic Press, Oxford, 2017).

Chapter 7. Summary and Outlook



In brief, this thesis explores the various applications of Raman spectroscopy and SERS in biosystem and biomedical applications. Using DCDR, the microenvironmental effects and the structural changes of the lysozyme protein, during aggregation were studied. SERS was combined with PCA to differentiate EVs extracted from various cell lines under starvation conditions. A PCR-based SERS strategy was developed to detect mutations in the tomato DNA. Finally, a compact tabletop Raman spectrometer was assembled and tested for future applications in bio-detection.

Many proteins aggregate under various conditions, including changes in temperature, pH and the presence of certain chemicals. Using the DCDR spectroscopy, the number of protein molecules under the laser spot increases, thus increasing the strength of the Raman signal. Various other proteins can be explored by this technique to study the microscopic effects on their structures and interactions. The microscopic effects of various small molecules can be explored to derive insights into their binding sites. Though protein aggregates polymerize into many structures, amyloid fibrils develop very specific structures. Their structures are being explored through many techniques, including x-ray diffraction, FT-IR and Raman spectroscopy. Raman spectroscopy can benefit from these other insights and prove very useful in detecting these fibrils. With better computational facilities and strategies, MD simulations can be used to mimic the realistic aggregation processes.

The classical extraction technique of EVs is the ultracentrifugation method, which in general is a long and extensive process. The use of a TEIR reagent hastens the process, allowing for quick extractions of the EVs. EVs extracted from a large number of cell lines need to be studied using SERS and PCA to obtain various clusters on the loading plots. The positions of these clusters need to be analysed along with the unique spectral features chosen for this plot to be able to differentiate the origin of these EVs. Further insights from biochemical analysis tools like proteomics, lipidomics, and genomics can help assign the Raman spectral bands, hence a better understanding of the molecular compositional analysis of these EVs by Raman spectroscopy. In the future, this technique can be explored as a diagnostic tool for studying EVs in live samples like blood or urine.

In this thesis, a PCR-based SERS strategy was developed to identify indels and SNPs in the DNA of a tomato plant. The same technique can be used for detecting many other kinds of DNA that have unique conserved regions, for example, the DNA of HIV. With the FWHM of the Raman bands being small, it is a very advantageous tool for multiplexing as compared to the many other techniques that exist in the

market today. By optimizing the oligonucleotide sequences, the type of metallic nanoparticles and the SERS conditions, this technique can be converted into technology for DNA detection.

Over the past many years, with the advancements in lasers, optics, electronics, and technology, there has been a tremendous improvement in the way Raman spectrometers have been built and miniaturized. With further improving the versatility of the system, the development of software to convert Raman spectral features into readable numbers, and the use of data analysis, Raman spectrometers can become devices in real-life scenarios rather than just being researched tools.

Overall, from this thesis, it is evident that Raman spectroscopy can be used in a wide variety of biosystem analysis and detection. The applications have a scope of being converted into actual technologies that can be useful to society.