EARLY AND RAPID DETECTION OF HUMAN IMMUNODEFICIENCY VIRUS IN BLOOD USING RAMAN SPECTROSCOPY

BEN OTIENO OTANGE

A Research Thesis Submitted to the Graduate School in Partial Fulfillment of the Requirement for the Award of Master of Science Degree in Physics of Egerton University

EGERTON UNIVERSITY

OCTOBER, 2017

DECLARATION AND RECOMMENDATION

DECLARATION

This research thesis is my original work and to the best of my knowledge, it has never been submitted in part or whole to any university or institution for any academic award.

Signature……………………………….. Date……………………

Ben Otieno Otange

SM13/23519/14

RECOMMENDATION

This thesis has been submitted for examination with our approval as university supervisors.

Signature……..……………………. Date ………………………
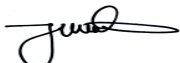Dr. Zephania Birech
Department of Physics
University of Nairobi

Signature………………………………Date ………………………
Dr. Ronald Rop
Department of Physics
Egerton University

Signature   [signature]   Date: October 12th. 2017

Dr. Julius Oyugi
Department of Medical Microbiology

University of Nairobi

ACKNOWLEDGEMENT

ABSTRACT

Deaths due to Human Immunodeficiency Virus (HIV) related diseases are still very high, especially in Africa. This is mainly attributed to poor diagnosis and late detection of the virus at times when the virus has greatly multiplied within the human body. In an attempt to come up with early detection mechanisms of the virus, several investigations have been carried out. The current detection mechanisms are less rapid with some being. Here we report on the potential of Raman spectroscopy as an alternative optical method for early and rapid detection of HIV-1 p24 in plasma and HIV-1 in whole human blood and in plasma. Raman spectra of HIV-1 p24 antigen, HIV-1 positive (HIV+) and HIV-1 negative (HIV-) blood and plasma were analyzed in the fingerprint spectral range 400-1800 cm^{-1} after 785 nm excitation. The Raman spectra of HIV1-p24 antigen displayed characteristic Raman bands centered at 891, 967, 1004, 1080, 1170, 1200, 1270, 1335, 1449, 1486, 1609, 1656 and 1738 cm^{-1}. Unique Raman bands were observed at around wavenumbers 928 cm^{-1} and 1658 cm^{-1} for HIV- and at 1270 cm^{-1} and 1446 cm^{-1} for HIV+ plasma samples. The prominent Raman band at 1270 cm^{-1} and 1446 cm^{-1} in HIV+ samples were ascribed to vibrational state of amide III of α-helix and C-H bond bending vibrational mode in proteins and lipids components of plasma respectively. For multivariate analysis; Principal Component Analysis (PCA) and Artificial Neural Network (ANN) were used. PCA was able to differentiate clearly Raman spectral data from HIV- and HIV+ plasma samples with sensitivity of 100% and specificity of 89.28% as well as sensitivity of 83.3% and specificity of 100% for whole blood samples. Classification based on ANN applied on the entire spectral range of data sets gave high correlation coefficient R = 0.99798 and R = 0.99997 for HIV+ blood and plasma samples respectively. Similarly, quantification also resulted in a linear model with $R^2 = 0.9457$ and $R^2 = 0.9959$ based on unique HIV+ peak intensity centered at 1270 cm^{-1} and 1446 cm^{-1} respectively. The high R^2 indicated the power of Raman Spectroscopy in detection and quantification of viral load both in blood and plasma. Each Raman spectral data was obtained within 15 s and thus the potential of being a rapid screening device.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF PLATES

LIST OF ABBREVIATIONS AND ACRONYMS

AIDS	Acquired Immunodeficiency Syndrome
ANN	Artificial Neural Network
ART	Antiretroviral Therapy
CCD	Charged Coupled Device
CE	Chemical enhancement
EDSDET	Electrochemical Detection Systems Based on Direct Electron Transfer
EDSTM	Electrical Detection System Based on Scanning Tunneling Microscopy
EDTA	Ethylene Diamine Tetra Acetic acid
ELISA	Enzyme-Linked Immunosorbent Assays
EME	Electromagnetic enhancement
HCA	Hierarchical Cluster Analysis
HIV-1	Human Immunodeficiency Virus type one
KNH	Kenyatta National Hospital
LSPR	Localized Surface Plasmon Resonance
MRS	Micro Raman Spectroscopy
NIR	Near Infrared
PBS	Phosphate-Buffered Saline
PCA	Principle Component Analysis
PCR	Polymerase Chain Reaction
PEP	Post Exposure Prophylaxis
PLSR	Partial Least Square Regression
PPB	Parts Per Billion
PPM	Parts Per Million
SERS	Surface Enhanced Raman Spectroscopy
UNITID	University of Nairobi Institute of Tropical and Infectious Diseases
WHO	World Health Organization

XRD   X-ray Diffraction

CHAPTER ONE

INTRODUCTION

1.1 Background Information

In the initial years, over 30 years ago, of acquired immunodeficiency syndrome (AIDS) history, human immunodeficiency virus (HIV) was unknown, misunderstood, untreatable, fatal, dreaded and often linked to traditions (Zaman *et al*., 2012). With the aim of understanding and eventually develop a cure for AIDS, HIV became the most intensively studied virus in human history (Coffin *et al*., 1996). Great progress has been made in obtaining an outline sketch of how genes and proteins in HIV particles operate and in understanding the biochemical specificity of HIV (Greene and Warner, 1993), the factors controlling its replication, the pathology of how it destroys the human immune system and the molecular bases of HIV infection and immunosuppression (Fauci and Anthony, 1993). There have been several attempts of developing new HIV detection techniques and most of them are yet to be available commercially. Some of the methods have aimed at detecting the HIV virus particles and they include Nano spectroscopic assays (Block *et al*., 2012), photonic crystal biosensors (Shafiee *et al*., 2014), electrical sensors (Shafiee *et al.,* 2015) and surface-enhanced Raman spectroscopy (SERS) (Lee-Ho *et al*., 2015 and Lee *et al*., 2015). Those that were designed to detect the HIV-1 p24 antigen includes electrochemical immunosensors (Ning Gan *et al*., 2013), electro-chemiluminescence immunosensor (Zhou *et al*, 2015), X-Ray Diffraction (XRD) methods (Raymond *et al.,* 2005) and Near Infrared (NIR) spectroscopy (Akikazu *et al*,. 2005; Rahim *et al.,* 2010).

Following the increasing HIV based research, there has been a corresponding development in detection methodologies over the years. For example, as at December 2016, 19.5 million out of the 36.7 million people living with HIV were under antiretroviral therapy ART. Comparatively, the number of infected persons under the same therapy has been on the rise especially in the last decade. For instance, the number was 15.8 million by June 2015 while only 7.5 million in 2010 and even much less in the previous years (World Health Organization, 2017). The turning point in the history of quest for cure of AIDS begun with the identification and linking of the HIV with AIDS. Thereafter, HIV diagnosis improved significantly with the development of highly effective Anti-Retroviral Therapy (ART) which enabled infected people to access treatment and lead a relatively normal and healthy life for a longer time.

There are two types of HIV which can infect humans; HIV-1 and HIV-2. Both types are transmitted in the same way and appear to cause clinically indistinguishable AIDS (Jensen, 2011 and Zaman *et al*., 2012). However, HIV-2 is less virulent and not easily transmitted. It is mainly found in western parts of Africa, and the period between initial infection and manifestation of illness is longer than for HIV-1. HIV-1 is the main cause of majority of the HIV infections globally. There are three subgroups of HIV-1; HIV-1-M, HIV-1-N and HIV-1-O of which HIV-1-M being the most common and widely spread globally (Madden *et al.,* 2011).

At the beginning of HIV infection, the virus antigen called p24 dominates and can be detected within 7-14 days after infection (Coffin, 1999). As the body develops a counteractive mechanism to fight the virus, HIV antibodies are produced about 90 days after initial exposure to the virus (Avert Group, 2015). The time taken to detect these antibodies may vary due to a number of factors but not earlier than three months. HIV infection gradually destroys the body's defense cells used in fighting infections, leaving the body susceptible to diseases it would normally be able to fight. When the virus is not diagnosed in time and proper monitoring mechanism is not put in place, the immune system becomes weaker, and the person with HIV will begin to develop opportunistic infections and hence manifesting AIDS (Zaman *et al.,* 2012 and Moor *et al*., 2013).

The widely used method for the detection of HIV in blood and recommended by World Health Organization (WHO) is based on the presence of specific antibodies associated with the virus (Ezzel, 2002). This method of diagnosis does not detect directly the presence of viral antigen or Ribonucleic acid (RNA) associated with the virus. Further, the method only works after about 90 days from initial exposure to the virus since specific antibodies are not produced before then (Moor *et al.,* 2013). Unfortunately, this is the only method accessed by the majority of diagnostic facilities in developing nations. At the early stage of HIV infection, the generic symptoms are often difficult to distinguish from those associated with common ailments such as influenza or fevers. Such symptoms include; sore throat, chronic diarrhoea, skin rashes, swollen glands, sweats especially at night and nausea (Zaman *et al.*, 2012). A number of studies have been done to formulate a method for early detection of HIV to enable a more effective prescription and treatment of secondary infections (Ghys, 2013). However, there are still shortcomings associated with the new methods in the market. These problems still need to be addressed through further research on early detection methods of HIV

Most of the HIV detection methods currently in use focus on detecting the virus itself rather than concentrating on the body's immune response (Avert Group, 2015). These methods include; Enzyme Linked Immunosorbent Assay (ELISA), Polymerase Chain Reaction (PCR), Optical and Nano technological methods such as Localized Surface Plasmon Resonance (LSPR), Detection System Based on Scanning Tunneling Microscopy (EDSTM) as well as Electrochemical Detection Systems Based on Direct Electron Transfer from Virus Particles (EDSDETV) (Lee *et al.*, 2015). These methods have various limitations in the detection process which include; low sensitivity, false positive results, high cost for large population and also labor intensive. These methods are summarized in table 1. Therefore, highly sensitive, selective, fast, and easy to use virus detection systems are needed for more effective diagnosis, accurate CD4/CD8 monitoring and for timely initiation of ART treatments.

Table 1: Summary of limitations of current early detection techniques of HIV-1

Method	Limitation(s)	References
ELISA	Monoclonal antibodies preferred here are very expensive and also difficult to get Requires a series of steps to finally get results False positive result in case of ineffective reagents	Juan Hu *et al*., 2010 Shafiee *et al.,* 2015
PCR	Labor intensive DNA amplification requires expensive reagents High rate of false positive and false negative result	Juan Hu *et al*., 2010 Akikazu *et al*., 2005
EDSTM	More time required before getting the result	Lee *et al.*, 2015
EDSDETV	Sensitivity is low in case enzyme is not used Current flow tempered by biomaterials affects sensitivity The method does not give accurate results in case of redox reaction	Lee *et al.,* 2015
LSPR	Cannot differentiate different binding events following multiple analytes	Lee *et al.,* 2015

Raman Spectroscopy is a vibrational spectroscopic technique used to collect a unique chemical fingerprint of molecules and hence identifying the molecules (Banwell *et al*., 2007). Viral infection causes some molecular and structural changes within the infected blood (Moor *et*

al., 2013). Specifically, the presence of p24 antigen (which is a protein) in a recently infected blood alters the molecular environment of the blood components (Moor *et al.,* 2013). Consequently Raman spectral signature of infected blood is expected to be different from that of healthy blood. The HIV-1capsid protein (HIV 1-p24 antigen) has great significance for diagnostics, because it can be detected several days earlier than host-generated HIV antibodies. The latter (HIV antibodies) are the target of majority of diagnostic tests used in the hospitals and laboratories. For early HIV diagnosis within the window period, sensitive assays specific to HIV-1 p24 antigen are needed (Ning Gan *et al*., 2013). Raman spectroscopic probe of biological tissues offers a good technique since it is sensitive to many different functional groups, offers highly selective fingerprints, requires no sample preparation, is compatible with aqueous solutions and can therefore be used easily to characterize the tissue and body fluids in nondestructive and noninvasive way (Smith *et al.,* 2005). Raman Spectroscopy is increasingly being explored as disease diagnostic tool due to its specificity, selectivity and also being label-free. In HIV detection, the bulk/classical Raman spectroscopy has exhibited its great strength in detecting HIV-1 antigen, HIV antibody and HIV antigen-antibody mixture (Pavel *et al*., 2011) and providing structural information of HIV-1 and HIV-2 in human sera (Chuanzong and Yiming, 2005). The other highly sensitive variant of Raman spectroscopy known as SERS has been used as an HIV gene probe (Narayana *et al*., 1998), HIV-1 DNA detector (Juan Hu *et al*., 2010) and HIV-1virus like particles detector (Lee Ho *et al*., 2015).

In this study, the potential of Raman spectroscopy in rapid detection of HIV-1 p24 antigen and HIV-1 in human whole blood and plasma was explored. The samples were aliquoted onto the exotic Raman substrate prepared by smearing of conductive silver paste onto a microscope glass slide. Unique Raman bands associated with HIV-1 p24 antigen were observed which were then used as biomakers during detection within the early stages after exposure. We report also the healthy and infected blood and plasma spectral characterization, segregation using chemometric tools and viral load quantification of the infected plasma samples (Otange *et al*., 2016; Otange *et al*, 2017).

1.2 Statement of the Problem

Early and rapid detection of HIV and subsequent intake of anti-retroviral drugs has been shown in many studies to lead to prolonged survival of infected persons. Current detection

methods such as ELISA, PCR, EDSTM, EDSDETV and optical methods such as LSPR suffer from myriad disadvantages. These shortcomings include; time interval required in detecting a specific immune response during the pre-seroconversion period and interpretation of latent infections, low detection capability, resource intensive or they are expensive to cover a wider population. The shortcomings have caused diagnostic dilemmas for the clinical virologist. In Kenya, majority of HIV diagnostic mechanism are designed to detect antibodies associated with the virus which become detectable at least three months after initial contact with the virus. In cases where concentration of antibodies in blood is low, detection of virus has never been successful. HIV cells may exist in the body of a person with severely impaired immune system who takes a relatively longer time to develop the specific antibodies. Similarly, persons who are exposed to HIV and administered with Post Exposure Prophylaxis (PEP) may take a longer time to develop these antibodies in case the drug fails to prevent the infection. For a new born baby, passive acquisition of maternal antibody may complicate confirmation of infection in the neonate. Therefore, this study demonstrates an alternative solution to some of the existing problems by providing a highly selective, sensitive, rapid detection and quantification of HIV-1.

1.3 Objectives

1.3.1 General Objective

To be able to detect rapidly HIV-1 p24 antigen and HIV-1 in human blood plasma and whole blood using Raman spectroscopy.

1.3.2 Specific Objectives

i) To prepare silver paste smeared glass slides for use as a sample substrate in Raman spectroscopy

ii) To determine the unique Raman spectral profile of HIV-1 p24 antigen, HIV-1 infected blood, HIV-1 uninfected blood and their corresponding plasma after laser excitation centered at 785 nm.

iii) To use Principal Component Analysis (PCA) and Artificial Neural Network (ANN) in the segregation of obtained combined spectral data from HIV- and HIV+.

iv) To use ANN and unique HIV-1 peaks of the obtained Raman spectral data from HIV+ samples for viral load prediction of the values obtained using PCR.

1.4 Hypotheses

i) Silver smeared glass slide have significant nanoparticle roughness necessary for the Raman signal enhancement.

ii) There exists a unique Raman spectral pattern of HIV-1 p24 antigen, HIV-1 infected blood, HIV-1 uninfected blood and their corresponding plasma.

iii) There is a significant difference between the spectral patterns of HIV-1 infected blood and uninfected blood as well as their corresponding plasma that could be used to classify similar Raman datasets in a common region based on PCA score plots and ANN training.

iv) There is a significant difference in Raman peak intensity related to different viral load that could be used to develop a linear quantitative model for HIV-1 viral load determination.

1.5 Justification

According to the UNAIDS 2016, by the year 2015, 36.7 million people globally were HIV infected, 2.1 million people became newly-infected and 1.1 million died of HIV-related causes. In comparison, WHO findings on HIV/AIDS prevalence worldwide in 2013, reported that 35 million people were infected and 1.5 million people died of HIV related causes that year. Sub-Saharan Africa accounted for 70 percent of all new infections and more than two-thirds of all deaths globally. Out of all the 35 million people living with HIV/AIDS, only 13.6 million people (about 40% of all people living with HIV) were receiving antiretroviral therapy (ART) as at June 2014. In addition, more than half of people living with HIV do not know their HIV status. In Kenya, National AIDS Control Council (NACC) in 2014 reported that 1.6 million people are HIV infected and HIV related deaths accounted for 29% and 15% national annual adults' and children' death respectively. Homabay County had the leading number of HIV infection with 25.7% of the total Kenyan living with HIV. However, the report found out that the county had less than expected people taking ARVs. Without effective HIV prevention, detection and or monitoring, the number of deaths as a result of HIV related illnesses may increase to alarming proportions. The current global HIV status therefore calls for effective, accurate and timely diagnosis of the virus in order to achieve the global goal of managing HIV and AIDS. Current methods being used for early detection of HIV-1 have serious limitations such as; detection selectivity, slow result acquisition and expensive reagents for covering a wider population. Raman spectroscopy is believed to have capacity to solve some of these challenges by providing a sensitive, timely and accurate method for rapid and potential early HIV-1

detection. Consequently, better diagnosis of the virus and hence significantly solve the current problems associated with late diagnosis and spread of the virus. The multivariate analysis techniques adopted in this study were based on their potential in retrieving important information from the patterns of the spectral data that could not ordinarily be noticed from the normal univariate analytical techniques and visual inspections.

CHAPTER TWO
LITERATURE REVIEW

2.1 Overview

Raman spectroscopy is a spectroscopic technique based on scattering of monochromatic light (Thermo Electron, 2003; Smith *et al*., 2005; and Kudelski, 2008). The scattering is such that the frequency of scattered radiations is either less or more than the incident radiation. It is an inelastic scattering in which frequency of photons in monochromatic light, usually a laser, changes upon interaction with a sample and the phenomenon is called Raman Effect (Smith *et al*., 2005). The change in the spectral frequency is called Raman shift and it provide unique information about vibrational, rotational and other low frequency transitions in molecules (Timothy, 2010). Each molecule has a unique vibrational energy level thus the photons emitted have a unique wavelength shift that are collected and examined to give the molecular identity (Pascut *et al.,* 2011 and Raquel *et al*., 2015).

2.2 Structure of HIV-1 virion

HIV-1 virion is approximately 120 nm in diameter and roughly spherical (Pavel *et al*., 2011). Mature virus consists of dense core containing the viral genome, two short strands of ribonucleic acid (RNA) about 9200 nucleotide bases along with the enzymes reverse transcriptase, protease, ribonuclease and integrase (Avert Group, 2015). All these are encased in an outer lipid envelope with 72 surface projections containing an antigen gp120 that aids in the binding of the virus to the target cells with CD4 receptors (see figure 1). The major structural components include the outer envelope glycoprotein gp120 and transmembrane glycoprotein gp41 derived from glycoprotein precursor gp160, nucleocapsid proteins p55, p40, p24 (core antigen), p17 (matrix), and p7 (nucleocapsid), enzyme proteins p66 and p51 (reverse transcriptase), p11 (protease), and p32 (integrase). Although most of the major HIV viral proteins, which include p24 (core antigen) and gp41 (envelope antigen), are highly immunogenic, the antibody responses vary according to the virus load and the immune competence of the host (Coffin, 1999). The antigenicity of these various components provides a means for detection of antibody, the basis for most HIV testing.

HIV has the additional ability to mutate easily, in large part due to the error rate of the reverse transcriptase enzyme, which introduces a mutation approximately once per 2000

8

incorporated nucleotides. This high mutation rate leads to the emergence of HIV variants within the infected person's cells that can resist immune attack, are more cytotoxic, can generate syncytia more readily, or can resist drug therapy (Ezzel, 2002).
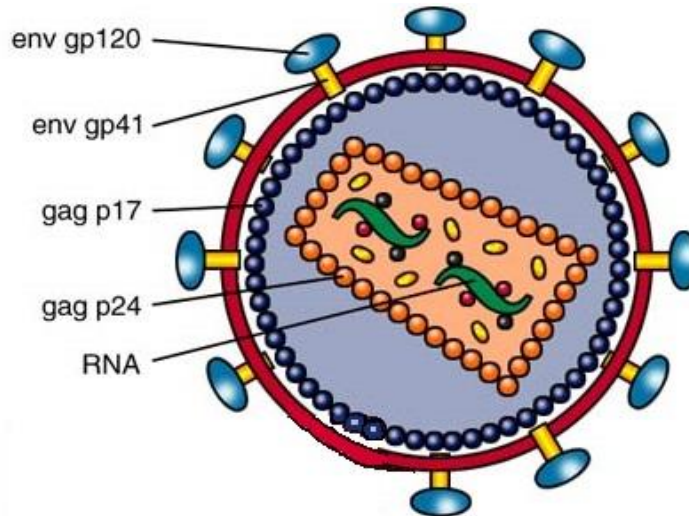
Figure 1: The Structure of HIV showing a viral envelope and viral core (adapted from Avert Group, 2015).

2.3 Structure of Proteins

The structure of a general amino acid which is the building block of proteins is shown below.

$$\overset{\overset{\textstyle R}{|}}{NH_2\text{-}CH\text{-}COOH}$$
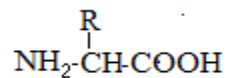
That means that the two simplest amino acids, glycine and alanine, would be shown as:

$$\overset{\overset{\textstyle H}{|}}{NH_2\text{-}CH\text{-}COOH}$$
glycine

$$\overset{\overset{\textstyle CH_3}{|}}{NH_2\text{-}CH\text{-}COOH}$$
alanine
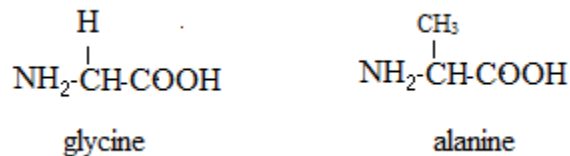
Glycine and alanine can combine together with the elimination of a molecule of water to produce a dipeptide. If three amino acids join together, a tripeptide is formed and if several amino acids are joined together (as in a protein chain), a polypeptide is formed. A protein chain will have somewhere in the range of 50-2000 amino acid residues (RCSB, 2011).
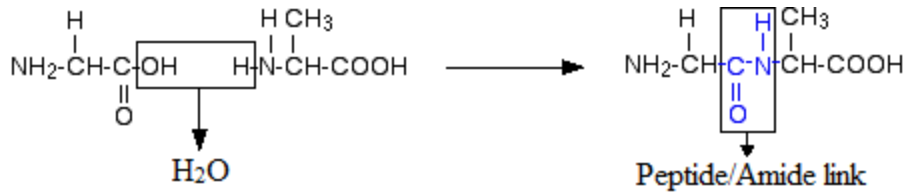
The peptide chain is made up from what is left after the water is lost. The end of the peptide chain with the -NH$_2$ group is known as the N-terminal, and the end with the -COOH group is the C-terminal. A protein chain will therefore look like this:
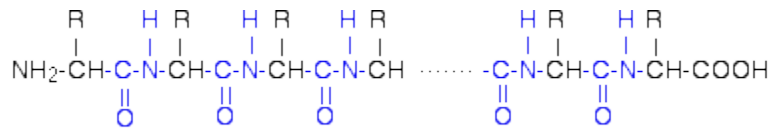
The "R" groups come from the 20 amino acids which occur in proteins. The peptide chain is known as the backbone, and the "R" groups are known as side chains. When all the elements are covalently bonded within the chain, the structure formed is called primary structure. For secondary structures (alpha-helices and beta-pleated sheets), the long protein chain are organized into regular structures whose side chains are held together by hydrogen bond. The tertiary structure of a protein is a description of the way the whole chain (including the secondary structures) folds itself into its final 3-dimensional shape as shown in figure 2.
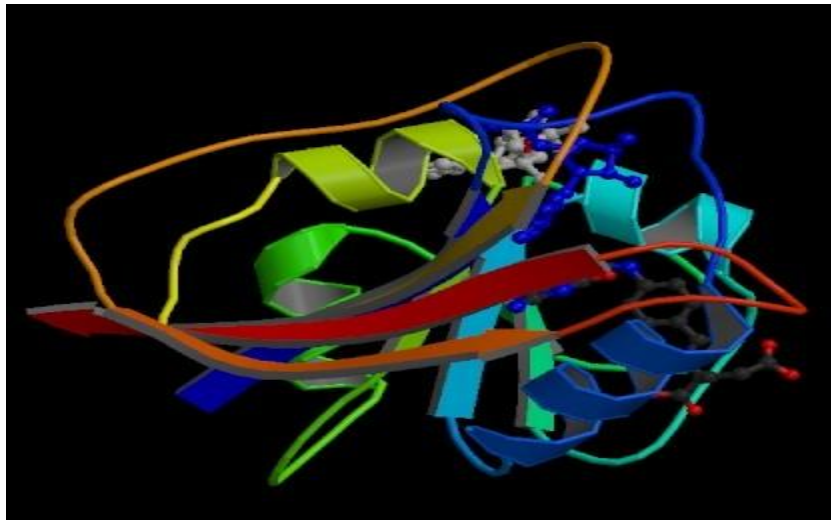
Figure 2: Final three dimensional shape of tertiary protein structure (adapted fromRCSB, 2011).

10

2.4 Theory of Raman Spectroscopy

2.4.1 Classical Theory of Raman Effect

When a molecule is placed within a static field, it experiences some charge distortion. The separation of charge centers sets in a dipole moment hence polarization. The size of induced polarization moment μ is dependent on the applied field E (Ying-Sing *et al.*, 2014).

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

Where μ is the induced dipole moment, α is the polarizability of the medium and E is the applied field given by;

$$E = E_0 \sin(2\pi\upsilon t) \text{ , } \upsilon \text{ is the frequency of the incident radiation}$$

$$\text{Therefore, } \mu = \alpha E_0 \sin(2\pi\upsilon t) \tag{2}$$

Since the ability to perturb the local electron cloud of a molecular structure depends on the relative location of the individual atoms, it follows that the polarizability is a function of the instantaneous position of constituent atoms. If polarization varies as the molecule vibrates with respect to some instantaneous position with coordinate q, then the new α is approximated by a Taylor series expansion;

$$\alpha = \alpha_0 + \frac{\delta\alpha}{\delta q}q \tag{3}$$

$$\text{Where } q = q_0 \sin(2\pi\upsilon' t) \tag{4}$$

q is the coordinate describing molecular vibration and υ' is the frequency of molecular vibration. Substituting q in equation (4) into equation (3) gives;

$$\alpha = \alpha_0 + \frac{\delta\alpha}{\delta q}q_0 \sin(2\pi\upsilon' t) \tag{5}$$

Substituting equation (5) into equation (2) gives;

$$\mu = \alpha_0 E_0 \sin(2\pi\upsilon t) + \frac{\delta\alpha}{\delta q} q_0 E_0 \sin(2\pi\upsilon t)\sin(2\pi\upsilon' t) \tag{6}$$

Applying trigonometry in the second term of the right hand side of equation (6) gives;

$$\mu = \alpha_0 E_0 \sin(2\pi\upsilon t) + \frac{1}{2}\frac{\delta\alpha}{\delta q} q_0 E_0 \{\cos 2\pi(\upsilon-\upsilon')t - \cos 2\pi(\upsilon+\upsilon')t\} \tag{7}$$

Thus the dipole moment oscillates with three different frequencies; υ, $\upsilon - \upsilon'$ and $\upsilon + \upsilon'$. The first term contains the variable similar to the frequency of the incoming light. This term relates to the outgoing scattered photon that has the same frequency as the incoming photon (Rayleigh scattering). The second and third terms contain the variable $\upsilon - \upsilon'$ and $\upsilon + \upsilon'$ which relates to an outgoing scattered photon that decreases and increases in frequency by some amount υ' which is the frequency of the molecular vibration. These two last terms ultimately constitute Raman scattering. The two terms show that incoming photons will shift their frequencies, down (Stokes scattering) and up (Anti-Stokes scattering) respectively, by amounts equal to frequency of molecular vibration (Banwell *et al.*, 2007, David, 2007)

From equation (7), it follows that the necessary and sufficient condition for Raman scattering is that the term $\dfrac{\delta\alpha}{\delta q}q_0$ must be non-zero for frequency shift to occur. This condition may be physically interpreted to mean that the vibrational displacement of atoms corresponding to a particular vibrational mode results in a change in the polarization (David, 2007).

2.4.2 Quantum Theory of Raman Effect

To discuss the inelastic scattering processes, a matrix element for the transition between initial and final state is evaluated. For the scattering process the transitions are driven by the polarization $P = \chi\varepsilon_0 E$ induced by the light. The matrix element, therefore, has the form

$$P_{fi} = \langle f \mid P \mid i \rangle = \langle f \mid \varepsilon_0 \chi E \mid i \rangle \tag{8}$$

Since wavelength of the light must be larger than the interatomic distance, the electric field can be considered to be constant in equation 8 so that a generalized form of the susceptibility known as the transition susceptibility can be extracted from the equation as shown.

$$[\chi_{mn}]_{fi} = \langle f \mid \chi_{mn} \mid i \rangle \tag{9}$$

Where χ_{mn} is a material specific quantity that is determined by the electronic orbitals in the crystal and if the final and initial states are both the ground state, it turns into the susceptibility. Applying adiabatic approximation (more details on adiabatic approximation see Chapter 7 of An Introduction to Solid State Spectroscopy by Kuzmany H, 2009) on equation 9 for electron and atom with coordinate x and X respectively and factored into wave function for the electron $\varphi(x, X)$ and for the atoms $\rho(X)$.

$$[\chi_{mn}]_{fi} = \int \rho_f *(X)\varphi_f *(x,X)\chi_{mn}\varphi_i(x,X)\rho_i(X)dxdX \tag{10}$$

Introducing normal coordinate Q_k, then expanding the electronic part of the susceptibility with respect to the normal coordinates while considering only the linear term of the expansion gives.

$$[\chi_{mn}]_{fi} = (\chi_{mn})_0 \langle ..v_{fk}..|..v_{ik}..\rangle + \sum_K \left(\frac{\partial \chi_{mn}}{\partial Q_k}\right)_0 \langle ..v_{fk}..|Q_k|..v_{ik}..\rangle \tag{11}$$

Where $\langle v_{fk}|$ and $|v_{ik}\rangle$ are the harmonic oscillator wave function for occupation numbers v_{fk} and v_{ik} respectively and the expectation value is given by.

$$\langle v_{fk}|v_{ik}\rangle = \begin{cases} 0 \text{ for } v_{fk} \neq v_{ik} \\ 1 \text{ for } v_{fk} = v_{ik} \end{cases} \tag{12}$$

and

$$\langle v_{fk}|Q_k|v_{ik}\rangle = \begin{cases} 0 & \text{for } v_{fk} \neq v_{ik} \\ \sqrt{\hbar/2\Omega k\left(v_{ik}+1\right)} \text{ for } v_{fk} = v_{ik}+1 \\ \sqrt{v_{ik}\hbar/2\Omega k} & \text{for } v_{fk} = v_{ik}-1 \end{cases} \tag{13}$$

As a result of orthogonality of the wave functions all the expectation values from equation 11 can be factored into equations 12 and 13. Then, the first term in equation 11 is only different from zero if $v_{fk} = v_{ik}$ for all k. This means, the quantum state of the system has not changed thus $(\chi_{mn})_0$ describes the process of Rayleigh scattering. The second term is responsible for the Raman process which is evident from the appearance of the derived susceptibility. According to equation 12 and 13 it is only nonzero if for all $k' \neq k$, $v_{fk'} = v_{ik'}$ and for the mode k, $v_{fk} = v_{ik} \mp 1$ holds. In this case the transition susceptibility in equation 11 has the following form

$$[\chi_{mn}]_{v_{ik+1},v_{ik}} = (v_{ik}+1)^{1/2}(\hbar/2\Omega_k)^{1/2}\left(\frac{\partial \chi_{mn}}{\partial Q_k}\right)_0 \tag{14}$$

$$[\chi_{mn}]_{v_{ik-1},v_{ik}} = (v_{ik})^{1/2}(\hbar/2\Omega_k)^{1/2}\left(\frac{\partial \chi_{mn}}{\partial Q_k}\right)_0 \tag{15}$$

Equation 14 and 15 describe the Stokes and the anti-Stokes Raman processes (Kuzmany H, 2009).

The incident laser photon is treated as a perturbation to the Eigen states of the molecule. This perturbation creates virtual vibrational states. When the incident photon excites the initial state, the photon stimulates the molecule to a higher level. This excitation is followed by a recombination with simultaneous emission of a photon of different energy. This state can decay back to the original state in such a way that no net change in energy is experienced (Rayleigh scattering). Depending on the status of the upper excited level where the molecule is excited to, the state can also decay to a state just above or below the original state resulting in an emitted photon that has either slightly more (Anti-Stokes Raman scattering) or less energy (Stokes Raman scattering) as shown in figure 3. This energy difference in the emitted photon is then due to the spread in energy level within the vibrational level and corresponds to the Raman shift (Banwell *et al*., 2007).
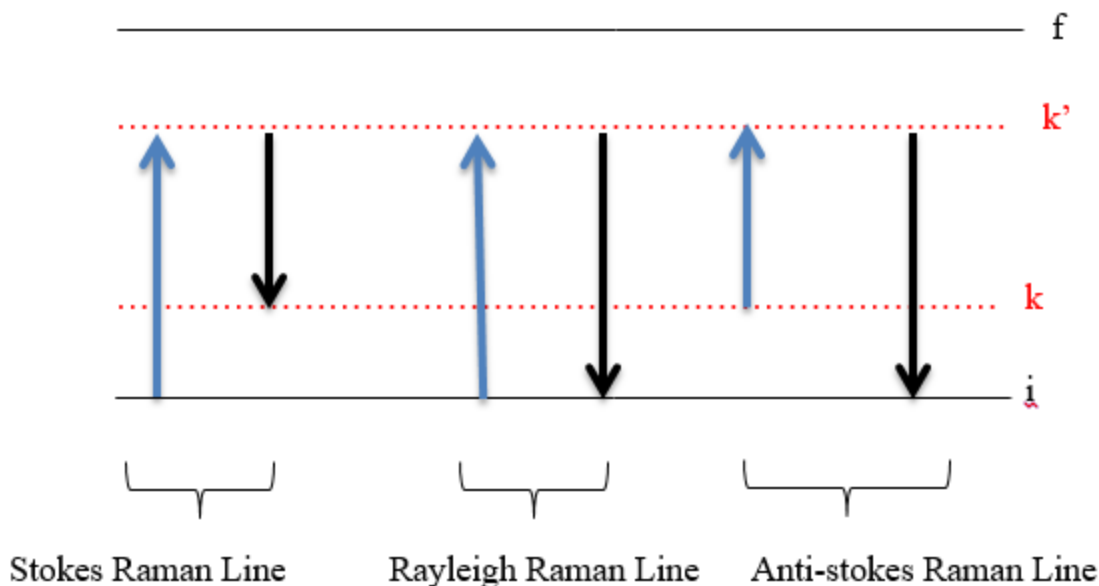
Figure 3: Energy level diagram for Raman scattering.

2.5 Micro Raman Spectroscopy

When Raman spectroscopy is combined with optical microscopy, spectral analysis of very small sections is made possible. This combination is called Micro Raman Spectroscopy (MRS) and is a powerful imaging technique with molecular specificity and lateral resolution capable of analyzing subcellular level. Coupling the strength and flexibility of Raman spectroscopy with a microscope allows analysis of very small samples. The main advantage of MRS in biomedical is its ability to provide functional imaging of live cells without applying labeling or sample

14

preparation (Moor *et al*., 2013). In order to get higher resolution, it is necessary to use smaller apertures. As light passes through these smaller apertures, diffraction which hinders proper image formation, becomes the limiting factor (David, 2007).

2.6 Surface Enhanced Raman Scattering

Surface Enhanced Raman Spectroscopy is a technique that extends the range of Raman applications to dilute samples and trace analysis, such as part per million level detection of a contaminant in water (Smith *et al.*, 2005). SERS shows promise in the fields of biochemistry, forensics science, food safety, threat detection, and medical diagnostics (Timothy, 2010). SERS technique provides greatly enhanced Raman signal from Raman-active samples that have been adsorbed onto certain specially prepared metal surfaces. SERS is achieved through two mechanisms of enhancement; an electromagnetic enhancement (EME) which is dependent on the presence of the metal surface's roughness features and a chemical enhancement (CE) which is dependent on changes of the adsorbate electronic states due to chemisorption of the analyte (Thermo Electron, 2003).

CE regards the enhancement of polarizability due to charge-transfer effect or chemical bond formation between the metal surface and molecules under observation and since this is associated with chemisorbed molecules, they can elucidate the mode of adsorption on the surface as well as the molecular orientation with respect to the surface plane. This bond produces surface species which include the analyte and some surface metal atoms that makes it possible to transfer electrons or holes from the metal surface into the analyte. The formation of this surface species is believed to increase the molecular polarizability of the molecule considerably due to interaction with the metal electrons (Chong-Shi, 2002).

The greater part of the overall enhancement of SERS is due to an EME mechanism which is the direct consequence of the presence of metal roughness features on the metal surface. EME takes into account interaction of the incident laser photon with irregularities on the metal surface such as metal micro-particles or roughness profile. The laser light excites conduction electrons at the metal surface leading to a surface plasma resonance and strong enhancement of electric field E (Thermo Electron, 2003). There are a number of ways of developing these features; for example: oxidation-reduction cycles on electrode surfaces, vapor deposition of metal particles onto substrate, metal spheroid assemblies produced via lithography, metal colloids and metal deposition over a deposition mask of polystyrene Nano spheres (Molly *et al*., 2010). All of these

methods leave the metal surface with small metal particles or aggregates of particles that are capable of serving as metal roughness features.

The two mechanisms simultaneously contribute to the total enhancement. SERS is observed primarily for analytes adsorbed onto the surface of Au, Ag or Cu as well as alkali metals such as Li, Na or K with the excitation wavelength near or in the visible region. The intensity of the Raman scattered radiation is proportional to the square of the magnitude of any electromagnetic fields incident on the analyte (Lin *et al.,* 2014). Smooth or flat surfaces are not good in an effective surface Plasmon generation. Surface roughness features with dimensions smaller than wavelength of the laser are preferred in SERS. Therefore, roughness features, or particle sizes in the range of 20-100 nm are commonly used for laser wavelengths in the range of 532-780 nm (Timothy, 2010). Both EME and CE theories agrees that the SERS enhancement requires nanoscale surface roughness. Using metallic colloids in aqueous solution is one of the ways of ensuring surface roughness. Colloids are quite easy to make, and provide a large surface area convenient for Raman studies in the laboratory (Smith *et al*., 2005).

2.7 Raman Spectroscopic Study of HIV

2.7.1 Bulk (Conventional) Raman Spectroscopic Study of HIV

A study was done to develop characterization spectra of HIV-1 p24 antigen, anti HIV-1 antibody and HIV-1 antigen-antibody complex using UV and NIR regions: 785 nm; 830 nm; and 244 nm laser excitations Raman spectroscopy by Pavel *et al.,* (2011). According to this study, Raman spectroscopy cannot be used as a tool to trace the reaction between the HIV-1 antigen and antibody, however, it can detect the HIV-1 p24 antigen (either independently or attached to the anti HIV-1 antibody). This is possible since Raman HIV-1 antigen peaks dominate the spectrum of antigen-antibody complex (Pavel *et al.,* 2011). The challenge arising from this study is that p24 antigen peaks at around 3 weeks after exposure after which its concentration drops significantly (Busch and Satten, 1997 and Stekler *et al.,* 2007). Therefore, this method is highly likely to give poor results when diagnosis is done during the seroconversion period when the p24 antigen level is very low.

Another study by Chuanzong and Yiming (2005), focused on structural damage of HIV-1 and HIV-2 in human serum and associated hypericin-induced photosensitive damage of the virus. Hypericin is used to prevent uninfected T-cells from being infected with the AIDS virus. Before then, it had been known that hypericin and some of polycyclic quinines irradiated by light

16

could inhibit HIV-1 replication; however, no study had given scientific proof of the same. This study provided direct evidence at the molecular level for photosensitive damage of HIV-1 and HIV-2 using bulk Raman spectroscopy. The resulting structural changes affected the binding sites of the virus thus hindering viral multiplication.

2.7.2 HIV-1 Detection based on Surface Enhanced Raman Spectroscopy (SERS)

Surface Enhanced Raman Spectroscopy (SERS) which encompasses the enhancement of potentially weak Raman signals has also been used in a sensitive, selective and rapid detection of HIV-1 as was reported by Lee-Ho *et al*., (2015). Their study focused on SERS detection of HIV-1 virus like particles (VPL) samples of different viral load based on immunoreactions with gold nano-dots fabricated indium tin oxide substrate without any labelling probe. Despite giving out a very good results that can be a potential immonoassay, this study did not employ blood, plasma or any other body fluid. According to this research, the higher the viral load of HIV-1 VLPs, the higher the Raman spectral spectral intensity. However, Raman spectral pattern of HIV-1 infection in blood and plasma show that peak intensity centered at 1446 cm^{-1} and 1270 cm^{-1} varies inversely with viral load (based on the findings of this research thesis).

Another HIV-SERS based study reported, investigated the combined use of PCR with SERS-labeled primers used in PCA reaction amplification of specific target DNA sequences (Narayana *et al*., 1998). This method combined the spectral selectivity and high sensitivity of the SERS technique with the inherent molecular specificity offered by DNA sequence of PCR. The effectiveness of this detection was demonstrated using the *gag* gene sequence of the HIV. The potential use of this detection probe, however, still depends on the PCR which is susceptible to a number of shortcomings (see background information).

A sub-attomolar HIV-1 DNA detection assay based on multilayer metal-molecule-metal nanojunctions was also developed using SERS (Juan Hu *et al*., 2010). Juan Hu and his coworkers developed a novel nanojunction based SERS biosensor with high sensitivity and selectivity towards DNA detection. They selected the HIV gene related DNA as a model due to its significance in early diagnosis and clinical therapy of HIV. Raman signal of the tag molecules on these detection probes was significantly enhanced. With regards to a HIV-1 DNA sequence, this SERS based biosensor platform could detect concentration as low as 10^{-23} moles thus providing a promising technique towards single molecular level HIV diagnosis. In as much as the assay gave out promising result for early detection of HIV, the cost of obtaining HIV DNA sequence, DNA

amplification and hybridization still challenges the method with regard to its applicability in a wider population.

2.8 Other Spectroscopic HIV Studies

Ever since the first instance of HIV/AIDS was identified, a number of studies have been focused on early detection, rapid result acquisition, improvement of sensitivity and selectivity of the virus detection method (Ghys, 2013). Some of the recent studies on the detection method have been linked to light wave interaction with molecules and the associated effect on the intensity, absorption, emission of the absorbed light and or scattering of light after interaction (Avert Group, 2015).

The applicability of NIR spectroscopy has been used in a series of studies in an attempt to come up with virus detection system. NIR spectroscopy has been successfully used to detect viral infections in fungal samples and further quantification of the viral infection using a multivariate classification and calibration model for the virus detected (Petisco *et al*., 2011). Plasma from HIV-1 infected individuals as well as the plasma of uninfected individuals (as a control) have also been subjected to NIR spectroscopy (Akikazu *et al*., 2005). The NIR spectra Tsenkova and his co-reserchers obtained from the two groups of samples revealed that NIR spectroscopy is also a promising technique in the detection of HIV-1 before the seroconvertion period following the confirmation by the reference ELISA method. In their work, they subjected NIR spectra in the range of 600-1000 nm to multivariate calibration model (PLSR) for further quantification of HIV-1 concentration within the infected plasma. Their analysis, even though never gave out an exellent prediction of the presence of the virus based on the values of the square of the correlation coefficient and standard error of cross validation values of the NIR spectra, it still provided an acceptable range of prediction based on the standards given by Petisco *et al*., (2011).

The combination of NIR spectroscopy with support vector machine has also been used to detect various types of HIV-1 infection. The NIR spetra in the region of 600-1100 nm was collected from HIV-1 molecular clone as well as the cultures of HIV-1 subtypes. This work revealed that the presence of HIV-1 in water affected the vibration of O-H bonds, thus affecting absorption at 950 nm and 1030 nm wavelengths. For different HIV-1 subtypes, the absorption was observed to be varrying at around 950 nm, 1030 nm and 1060 nm suggesting that different

subtypes of HIV-1 present in water affect the O-H vibration differently thus facilitating their detection (Rahim *et al*., 2010).

2.9 HIV-1 Detection based on Nanotechnology

There has been a number of suggestions on detection methods of HIV-1 based on nanotechnology such as EDSTM, EDSDET and LSPR methods. EDSTM consists of voltage applied between a sharp metal probe and an electrically conductive material that are separated by about 10 nm in such a way that a very small tunneling current is produced between the terminals. This tunneling electrical current passes through the neighboring media (conducting surface) and is scanned by the metal tips. A change in tunneling current proportional to different density states in accordance to quantum mechanics can be used to provide a real current profile of the conducting media. The presence of HIV-1 within the conducting media will give a different current profile thus can be a new method of detection of the virus (Lee *et al*., 2015).

Based on optical LSPR and nanotechnology, HIV-1 can also be detected following the changes in refractive index induced by the adsorbate on the metal nanostructure which provides information that can be used to study molecular binding events. These changes in the refractive index on the metal surface vary the Plasmon resonance conditions at different frequencies, wavelength or angle shifts depending on the adsorbate medium (Lee *et al*., 2015) They demonstrated the usage of LSPR in detecting HIV by measuring the absorbance changes on gold nanostructure surface and found out that the presence of HIV-1 within the adsorbate increased the absorption peak.

Based on the challenges still facing the early HIV-1 detection methods currently being used, there is a need for further research in this field to solve some of these existing problems. These challenges include; selectivity of the method in virus detection, rapid result acquisition, expensive reagents as well as simplicity in covering a wider population. Raman spectroscopy is believed to have capacity to provide a solution to some of such challenges.

2.10 Multivariate chemometric techniques

Chemometrics techniques refers to a chemical discipline that utilizes statistical and mathematical methods to design or select optimum procedures and experiments as well as to provide maximum extraction of relevant information by analyzing spectroscopic data (Hanson, 2015). Multivariate analysis uses several variables at the same time by considering main variance

expressed as meaningful interpretable factors (loadings). Multiple algorithms are available with chemometric techniques and the choice of the technique depends on the desired interest from the data output. For decomposition of data output Principal Component Analysis (PCA), Multiple Curve Regression (MCR) may be used. Classification algorithm include; PCA, Hierarchical Cluster Analysis (HCA), Discriminant Cluster Analysis (DCA), Linear Discriminant Analysis (LDA) and Artificial Neural Network (ANN). Regression analysis for quantification purposes is performed by Partial Least Square Regression (PLSR), Principal Component Regression and ANN.

In this research, focus was on the multivariate Raman spectral datasets from different samples. The task was to retrieve important information from the patterns of the spectral data that could not ordinarily be noticed from the normal univariate analytical techniques and visual inspections. This reseach work combined the usage of PCA and ANN to give a more effective Raman-Chemometric algorithm for detection of HIV-1 infection through spectral classification.

2.10.1 Principal Component Analysis (PCA)

Principal Components Analysis PCA is a useful statistical technique applied in fields such as face recognition and image compression, and is a common technique for finding patterns in data of high dimension as well as expressing the data in such a way as to highlight their similarities and differences (Marleen *et al.*, 2006). PCA is one of the main methods of identifying dominant clusters in datasets by use of eigenvalue-eigenvector technique of matrix algebra. PCA decomposes correlation of covariant matrix into eigenvalues and corresponding eigenvectors which are orthogonal to each other. The main advantage of PCA is that pattern in the data can be obtained and compressed by reducing the number of dimensions, without much loss of information (Smith Lindsay, 2002). PCA does this by sorting out the most influential factors (principal components) from the data set followed by less influential factors, compressing the dataset by keeping only these important information, simplifying the description of the dataset and analyzing the structure of the observations and the variables. As a result, unwanted information is eliminated without much loss of information (Herve and Lynne, 2010).

A number of tasks in medical and pharmaceutical fields have employed PCA for various reasons including; ascertaining the spectral variance within the population of cells and detecting underlying patterns within the data from the chloroquine treatment group and the control group (Grant *et al.*, 2008), designing an effective algorithm to discriminate the SERS spectra between

the healthy and colorectal cancer serum samples (Lin *et al.,* 2011), building a chemometric classification model for the adenovirus infected cells (Moor *et al*., 2013) and a fast detection and identification of counterfeit antimalarial tablets Raman dataset (Marleen *et al.*, 2006)

2.10.2 Artificial Neural Analysis (ANN)

Artificial neural networks are nonlinear computational tools capable of modelling complex functions such as Raman spectral data which at times are full of noise. ANN is useful in non-linear ordination and pattern matching algorithm that results in classification, regression and or clustering of tasks (Ajith Abraham, 2005 and Hanson, 2015). ANN is a non parametric modelling tool in form of computation units arranged in layers mimicking the physiologic structure of the brain to learn from the data using a series of weights and hidden neurons to detect complex relationships (Grossberg, 1976). There are multiple connections between units within and between layers. These connections have strengths or "weights" that are "learned" by the network. The training paradigm is either "supervised," where sample input-output pairs are presented (Rosenblatt, 1959) or "unsupervised," where the network organizes itself (Kohonen T, 1984). This "learned" information in the network is stored in these interconnection weights.

A variety of medical tasks that have been successfully performed using ANN includes; lesion detection in SPECT images (Floyd and Tourassi, 1992), prediction of pulmonary embolism from perfusion scans (Scott and Palmer, 1993), breast cancer analysis (Dawson *et al*., 1991) and decision-making in mammography (Wu *et al*,. 1991 and Wu *et al*., 1993).

2.11 Univariate Raman Spectral Analysis

Raman spectra provide a unique fingerprint which represents the set of bonds present in the sample. The spectral signature produced is a function of vibrational frequencies of the chemical bond or group of bonds within the sample (Jason *et al.,* 2013). Four pieces of information are very important in analyzing these spectral signatures; they include: band position, frequency shift, line width and intensity as shown in figure 4. Vibrational frequencies are sensitive to details of structure and local environment of a molecule, such as symmetry, crystal phase, polymer morphology, solvents and interactions (Kudelski, 2008). Qualitative Raman Spectroscopy is concerned with identification of components through analysis of band position (David, 2012). Band position provides information on chemical species, crystal phases and alloy composition (Dinh *et al*., 2005). Band position may shift due to changes in environmental factors

such as temperature and strain which may bring about a frequency shift without changing the spectral pattern (David, 2012). Frequency shift provide information on molecular interaction with the environmental factors and chemical reactions
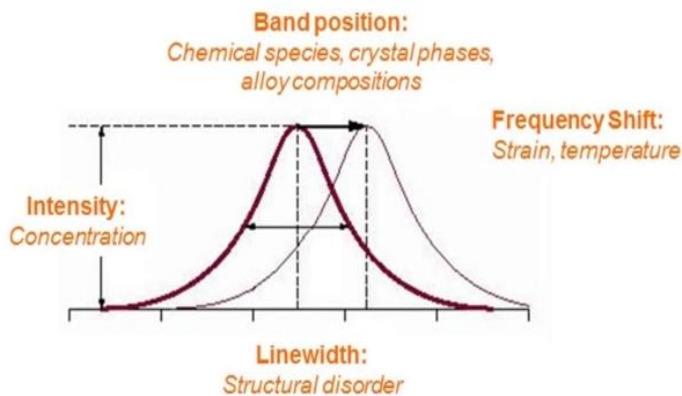
Figure 4: A schematic of a Raman spectral band showing the various parameters including band position, line width, intensity and frequency shift (adapted from David, 2012)
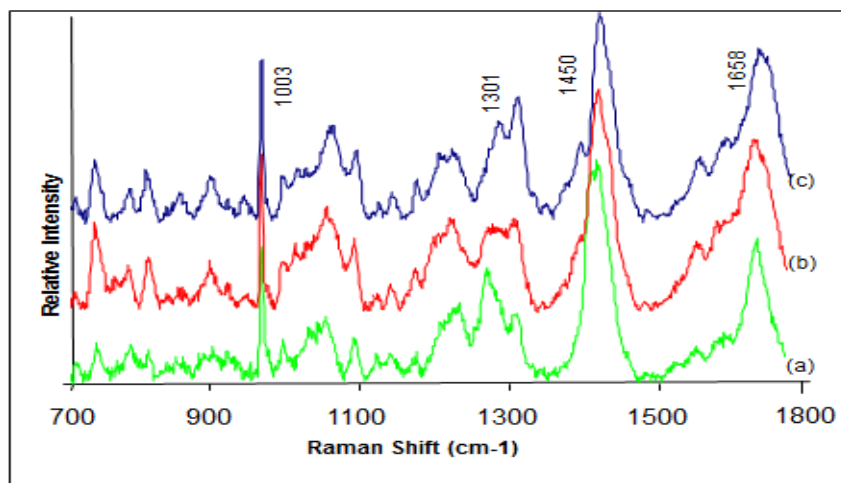
Figure 5:   Quantitative analysis based on relative intensity; Control spectra (c), 24 h after infection (b) and 7 days after infection (a) (adapted from Moor *et al.,* 2013)

Quantitative analysis of a sample is performed by measuring the relative intensities of bands that are directly proportional to the relative concentrations of the various components within the sample as shown in figure 5. Alternatively, Chemometric methods can be used to give quantitative analysis when performed on samples with high concentrations ranging from 90-

100% material of interest down to concentration determination at parts per billion (PPB) and parts per million (PPM) levels (Smith *et al*., 2005).

2.12 Application of Raman Spectroscopy in medical and pharmaceutical fields

Vibrational Raman spectroscopy provides key information on the structure of molecules since no two molecules can give exactly the same Raman spectrum. Different molecules have distinct vibrations, thus the Raman spectra of molecules are unique making identification easier. The position and intensity of features in the Raman shift spectrum is used to study molecular structure and to determine the chemical identity of the sample (Thermo Electron, 2003).

Raman spectroscopy has been explored as a fast and reliable screening method for the detection of counterfeit tablets in various pharmaceutical areas as well as sensitive, selective and real time detection of various infections (Gelder *et al*., 2007). For instance, this technique has been used to detect and identify the components of anti-malarial counterfeit tablets from the difference in their spectral pattern as compared to that of the genuine anti-malarial tablet (Marleen *et al*., 2006).

Recently SERS was employed in detecting blood and plasma fingerprints for gastric cancer (Feng *et al.*,2011), medical diagnosis and biological imaging (Dinh *et al*., 2005), colorectal cancer (Lin *et al.,* 2011), cervical cancer (Feng *et al.,* 2013) and nasopharyngeal cancer (Feng *et al*., 2010; Lin *et al*., 2012 and Lin *et al*., 2014), dental prosthesis (Gyeong Bok *et al*., 2014). Recent studies on biological infections not only tend to focus attention on early detection of the infection but also examine the effect of the infection on body components for better diagnosis. The capability of SERS to discriminate between blood and plasma components from healthy volunteers and patients based on individual blood and plasma components such as DNA/RNA, lipids, carbohydrates and proteins have also been demonstrated with well defined tentative peak assignment (Feng *et al.,*2011; Lin *et al*., 2014; Feng *et al.,* 2013 and Lin *et al*., 2014).

Apart from identification of counterfeit drug tablets and detection of various infections, Raman spectroscopy can also be applied in the analysis of drug-cell interaction in order to probe not only the presence of the drug in the cell but also the effects of the drug on the cells (Grant *et al*., 2008). In recent years, Raman scattering-based spectroscopy and microscopy have been extensively used to determine the effects of chemotherapeutic agents on living eukaryotic cells, as well as antibiotics and other small molecules in prokaryotic cells. Raman spectroscopy in this

regard, is useful because it enables one to follow the long-term effects of drug treatments on cells, potentially identifying cells that avoid apoptosis and determining the minimum dose needed to avoid minimal residual disease, a particular problem in the fight against the recurrence of cancer (Anupam *et al.,* 2011). A similar study was applied to monitor the effects of chloroquine treatment on cultures of *Plasmodium falciparum* trophozoites. The spectral result showed that intensity changes are attributed to intermolecular drug binding of the chloroquine in a sandwich thus demonstrating the potential of Raman microscopy as a screening tool for drug and live cells interaction (Grant, 2008).

CHAPTER THREE
MATERIALS AND METHODS

3.1 Sample collection

Blood and its corresponding plasma samples were obtained from 30 volunteers (22 patients with confirmed clinical diagnosis of HIV-1 and 8 healthy individuals) after a consented agreement (see Appendix A, B and C). The collected samples were then anonymized by assigning them to unique identifiers, thus eliminating chances of tracing them back to the specific volunteer in accordance to the confidential agreement. All volunteers provided written informed consent to participate. This study was approved by the Kenyatta National Hospital-University of Nairobi (KNH-UoN) Ethics and Research Committee (Proposal number: P637/10/2015 - see Appendix D). Samples were collected after overnight fasting in Ethylene Diamine Tetra Acetic acid (EDTA) collection tubes and temporarily stored in a vertical position (see Appendix F Plate 1). Blood and plasma were then separated based on their densities, with plasma (a pale yellow fluid) occupying the upper part being less dense (see Appendix F Plate 3). The plasma was then pipetted out into a different EDTA container. After separation, PCR test was performed on the plasma samples to classify them as either HIV-1 positive or negative samples and for positive plasma to determine their viral load.

3.2 Safety and Sample Storage

The blood analytes, plasma and P24 antigen suspended in Roswell Park Memorial Institute (RPMI) 1640 medium and Phosphate Buffered Saline (PBS) were stored at $-20^{o}C$ within unopened ampoule in the UNITID laboratory. From the storage site, the samples were transported directly to the Laser lab for spectral fingerprinting. P24 antigen, positive blood and plasma samples were regarded as potentially hazardous to health and were therefore used and discarded according to laboratory's safety procedures (see Appendix E). Before analysis of p24 antigen, the sample was placed in a hot water bath at $56^{0}C$ for one hour within the unopened ampoule to deactivate it and reduce chances of infection should it come in contact with a cut.

3.3 Calibration of Raman Equipment

The Raman equipment was calibrated daily before taking measurement using the 520.5 cm^{-1} band of a silicon wafer. This calibration was done to check possible Raman peak shift that

may occur as a result of temperature. During this process, the 520.5 cm^{-1} peak wavelength shift was adjusted manually.

3.4 Raman substrate preparation

The glass slides were cleaned using ethanol and then air dried for about 20 minutes. The silver conductive paste/paint (SPI suppliers, USA) was smeared using a small brush onto a microscope glass slide and air dried for an hour. The silver paste consisted of silver metal particles (35-65% of total weight), 1-methoxy-2 propanol acetate (10-30% weight), butyl acetate (10-30% weight) and acrylic resin (5-10% weight). The procedure was repeated till the silver colloids and silver deposition over a deposition surface appeared as nanospheres. This preparation left the glass slide surface with nanoparticles or aggregates of particles that are capable of serving as metal roughness features as shown in Figure 6a and Figure 6b under x10 objective lens (owing to its spectral lower signal to noise ratio as compared to other objective lenses). On different Raman substrate prepared, p24 antigen, RPMI, PBS, blood and plasma samples were smeared on the surface and the adsorbate left to dry for about one hour 30 minutes before taking Raman measurements.
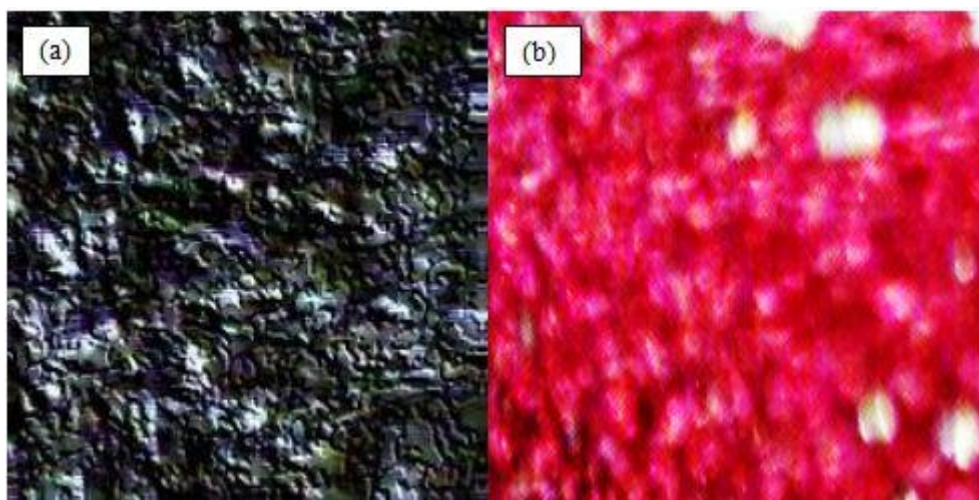
Figure 6: Ag substrate (a) and blood adsorbed on Ag substrate (b)

3.5 Raman Optimization Parameters

Raman spectra were obtained using a Laser Confocal Raman microscope (STR Raman Spectrum System, Seki Technotron Corp; Model number RO -110J, see Appendix F Plate 5 and 6) with laser excitation wavelength of 785 nm. The laser was focused on the samples with x10 objective lens (Max Plan). Raman scattered light from the sample was collected using the same

26

objective lens and detected by a Charge Coupled Device CCD (Princeton Instruments; Acton SP2300) equipped with a 256 x 1024 pixel camera cooled at -76^0 C. The excitation parameters were; diameter of laser spot at the focus point $\approx$ 71µm, excitation power was 150 mW, exposure time of 15 s and 15 accumulations per spectra. The grating chosen in order to cover a wider spectral range was 600 lines per mm grating. Background spectra were collected while blocking the laser source from being scattered by the sample. For each samples, spectra were recorded from 10 different spot areas. Raman spectra were then analyzed in the 'finger print' spectral region 400-1800 cm^{-1} with center wave number of 1100 cm^{-1}.

3.6 Spectral Fingerprinting

The schematic diagram of Raman spectroscopy set up is shown in figure 7. Dispersive MRS was carried out using system that coupled optical microscopes with conventional Raman spectrometers. The incorporation of a microscope with a three dimensional stage enabled specific regions of a sample to be analyzed.
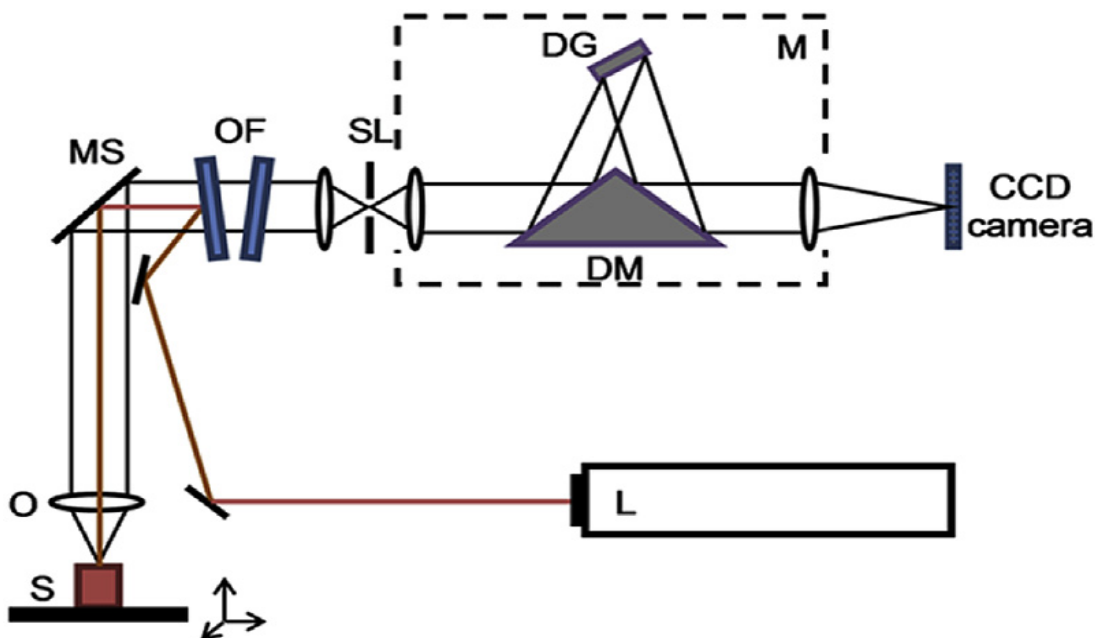
Figure 7: A Schematic diagram showing dispersive Raman set up (Ying-Sing *et al*., 2014)

The laser (L) was focused through the microscope (MS) onto the sample (S) and the Raman scattered light collected by the same microscope objective (O). The scattered radiation was passed through optical filters (OF) to remove the Rayleigh line prior to being directed through slits (SL) and into the monochromator (M) consisting of diffraction grating (DG) and a

27

dove mirror (DM). The removal of Rayleigh scattered lines by optical filter is the only remedy for obtaining good Raman spectrum since Rayleigh lines accounts for about 99.999%. Finally the Raman scattered radiation was focused onto the CCD camera to reveal the spectral pattern.

The samples were placed under the microscope and positioned visually using the CCD camera and three dimensional translation stage. A photograph of the tissue was then taken and after which the light and camera were turned off and the camera's mirror removed. The laser beam was then turned on. The spectra taken were:

i) Raman spectral profile of background radiation

ii) Raman spectral profile of silver substrate colloids on the glass slide.

iii) Raman spectral profile of smears of p24 antigen suspended separately on PBS and RPMI-1640 and smeared on the silver substrate.

iv) Raman spectral profile of the smears of salt preservatives; PBS and RPMI-1640 on the silver substrate.

v) Raman spectral profile of plasma adsorbed on silver substrate

vi) Raman spectral profile of plasma contaminated with p24 antigen adsorbed on silver substrate

vii) Raman spectral profile of blood adsorbed on silver substrate

3.7 Data Collection

The data obtained from each sample in each of the stages comprising of Raman shift in wave number units and the corresponding intensity of the peaks in arbitrary units were saved within the Raman computer system based on date of sample collection, sample(s) used, and optimization parameters so as to avoid confusion in the voluminous data that was collected at any single exposure time. From the Raman computer, the data was then transferred to another computer for chemometric analysis.

3.8 Data pre-processing and analysis,

Further data pre-treatment were done using three softwares: Origin Professional 2015, Vancouver Raman Algorithm and MATLAB R2009a release. Data pre-treatment is a set of transformations that modify the data making the data to be tractable for direct use. Samples were first smoothened using Savitzky-Golay filtering function (at 9 points, second derivative) in MATLAB. Spectral smoothing and filtering was done to denoise the Raman dataset. Second

derivative transformation was used to enhance the separation of overlapping band components hence making spectral comparisons much easier (Persson, 2003). Savitzky-Golay smoothing method was preferred as it gives a high degree of polynomial fit, provide computational time proportional to window width coupled with the advantage of preserving the area, position and width of the peaks useful for some further analysis (Savitzky Golay, 1964 and Persson, 2003). Vancouver Raman Algorithm program was used to extract Raman signals from the measured raw data that usually consists of the pure Raman signals superimposed on background auto fluorescence signals (JIANHUA *et al.,* 2007). Origin software was used for plotting and spectral subtraction to isolate the unique fingerprints of p24 antigen, blood and plasma from silver substrate. The number of peaks and their approximate positions were first identified from the minima observed in second derivative of combined spectra. The spectral region of interest was then modeled using a corresponding series of initial band shapes after which the peak maximum frequency, peak height, width, and shape are then allowed to vary till a best fit was obtained.

3.9 Chemometrics

The use of chemometrics in relevance to multivariate data analysis provide quicker and better pattern recognition, classification, linear and nonlinear mapping of the data. Spectroscopic methods are highly multivariate in nature and require powerful visual and analytical tools to unlock their rich information. Multivariate data offers a multidimensional data measured on a set of similar samples, cases, objects and or process points. The simple principles of models, multidimensional spaces, and projections provide a strategy for utilizing this richness of information (Wold, 1995). Chemometric technique provides a more realistic hypothetical description of the multidimensional data analysis than the often misleadingly exact looking relationships. This study particularly used PCA and ANN since they all can take into consideration the joint effect of all variables unlike the physical difference between variables which often considers only one variable to analyze data. The physical difference as a traditional technique cannot therefore be used to analyze latent variables within a multivariate data.

3.9.1 Principal Component Analysis (PCA)

Principal Component Analysis - was used to develop a classification model that could discriminate between the infected and healthy whole blood as well as the blood plasma Raman datasets. This was achieved by highlighting all the dataset of a given fluid type for both the

infected and control samples and subjecting the dataset to default PCA dialog to generate a report sheet. From the generated report sheet, the eigenvalues of the correlation matrix table displayed the number of principle components (PCs) to be considered based on the percentage of the principal components with respect to the total variance. Similarly, from the scree plot within the report sheet, the elbow point could also tell the number of PCs to be selected. Having chosen the PCs with the higher percentage of total variance, a PCA dialog was reopened to input new parameters and recalculate based on the chosen parameters. The resulting loading plots could discriminate the HIV- and HIV+ spectra with high sensitivity and specificity. Each of the spectral data was plotted as a score on a set of two PCs (score plot). Each PC is associated with a spectrum loading vector and its loadings plot has the appearance of a spectrum (Alexander and Tapani, 2010; Herve and Lynne, 2010).

3.9.2 Artificial Neural Network (ANN)

Artificial neural networks are nonlinear computational tools capable of modelling complex functions such as Raman spectral data which at times are full of noise. ANN is useful in non-linear ordination and pattern matching algorithm that results in classification, regression and or clustering of tasks (Ajith, 2005 and Hanson, 2015). ANN is a non parametric modelling tool which mirrors biological neurons to learn from the data using a series of weights and hidden neurons to detect complex relationships (Grossberg, 1976). In order to achieve a better ANN analysis, the network was exposed to a set of voluminous data for self organization that leads to discovery of pattern.

The developed ANN used Raman spectral dataset as inputs and predicts the outcome of training as an output that was then applied in classifying the target dataset as either HIV-1 positive or negative. This "trained knowledge" was contained in the internal numeric "weights" of the network. The training algorithm repeatedly modified the numeric values of these weights to decrease the training error of the network. Once trained, the weights were fixed and assimilated for reproducible results. When a set of input feature values were presented to the trained network, an output was generated that represents the classification of the prediction. This output was based on the trained knowledge that the network learned in the training step.

The network was trained and evaluated on all cases using the jack-knife technique. This technique allows utilization of all available cases for training but also provides a meaningful evaluation of the generalizing ability of the trained network. With this technique, one case was

30

selected for evaluation and the network was trained on the remaining cases. The trained network was then evaluated on the single case that was left out. The selected case was put back in and another case was removed. This process was repeated until all cases had been used for evaluation.

3.10 Quantification based on univariate Raman peak analysis

Selected prominent peaks believed to be associated with HIV-1 infection in plasma were used to quantify viral load. Plasma from healthy donors and individuals infected with HIV-1 were subjected to HIV-1 PCR test and Raman spectroscopy. After the collection of Raman spectra, the spectral data were pre-processed and subjected to peak calibration modeling to develop a univariate model of estimating the concentration of HIV-1 in plasma. The model was confirmed by a leave-out cross validation. When the model was not acceptable, re-calibration was performed after the validation as shown in the flow chat in figure 8.
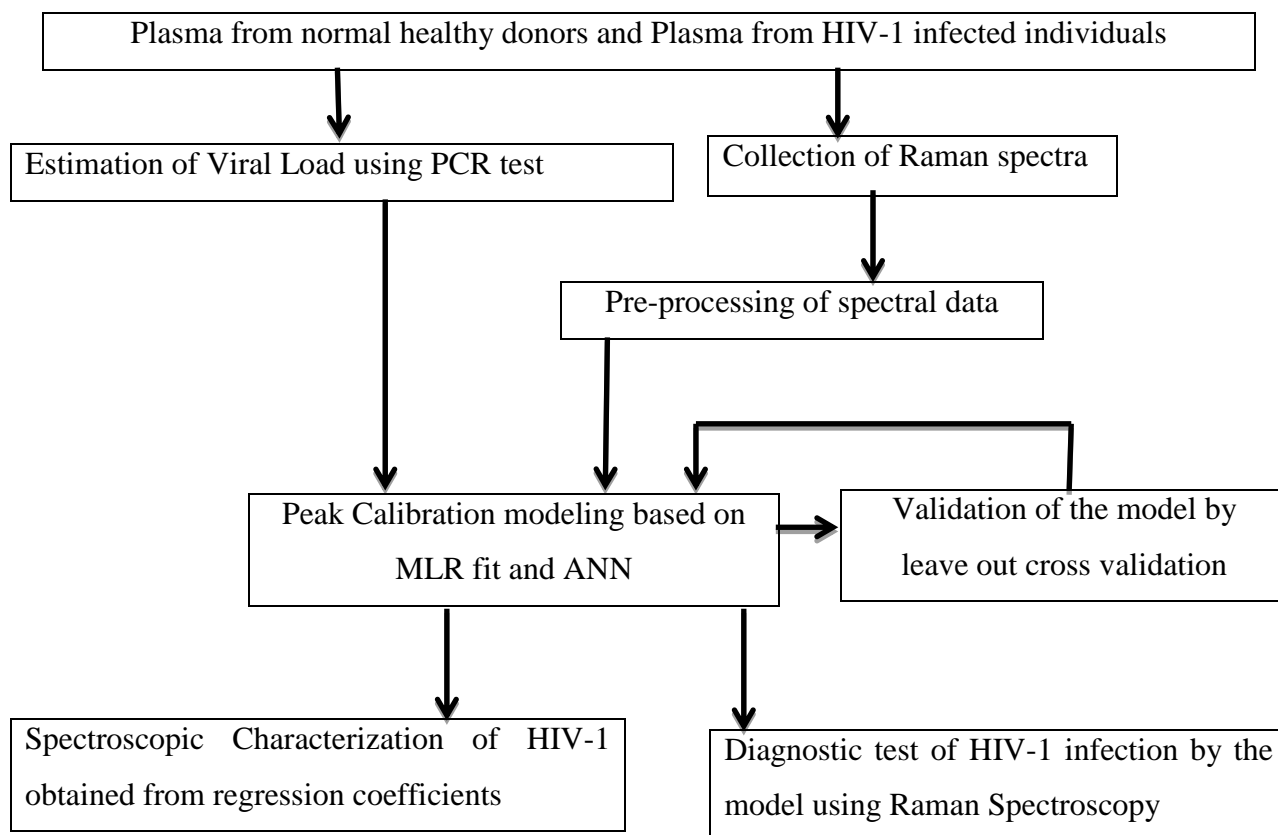
Figure 8: SERS calibration flowchart of HIV-1 load estimation

CHAPTER FOUR
RESULTS AND DISCUSSION

The results from the study findings are summarized in this chapter with special attention on the verification of the study objectives. We indeed verified that the silver nanoparticles on the microscope glass slide provided great p24 antigen Raman signal enhancement that enabled its characterization. This was thought to have been caused by attachment of HIV virion to the silver nanoparticles via the gp 120 glycoprotein knobs. This chapter reported on the characterization of; HIV-1 p24 antigen, HIV- and HIV+ whole blood and plasma, Segregation of healthy and HIV-1 infected spectral dataset for whole blood and plasma based on PCA and ANN as well as the quantification of viral load amongst the infected samples with reference to PCR. Parts of the findings in this section were published and appear in the papers by Otange *et al*., (2016) and Otange *et al*., (2017).

4.1 Raman Spectroscopic Detection of HIV1-p24 antigen

In order to be able to detect HIV-1 p24 antigen via Raman spectroscopic techniques, one has to have idea of its characteristic Raman spectrum. Here we show how we were able to obtain this special characteristic features which are rare in literature. Having known the characteristic spectra of HIV-1 p24 antigen, its presence within intentionally contaminated plasma was sort with intentions of demonstrating early detection spectroscopic methodology.

4.1.1 Characteristic Raman spectrum of the HIV-1P24 antigen

HIV-1 p24 antigen is situated at the core of HIV-1 virus particle (HIV virion) and I composed of two strands of viral RNA genome and proteins; nucleocapsid p7, reverse transcriptase p66/p51 and integrase p31 (). When the HIV virion is excited, it is therefore expected that the spectral profile will have bands associated with components constituting it. The results showed that the characteristic spectra were indeed associated with RNA and proteins. Vancouver Raman Algorithm based on fifth order polynomial fitting method developed by Zhao *et al.,* (2007) was used to subtract the auto fluorescence background Raman signal superimposed on the desired Raman signals as shown in Figure 9.
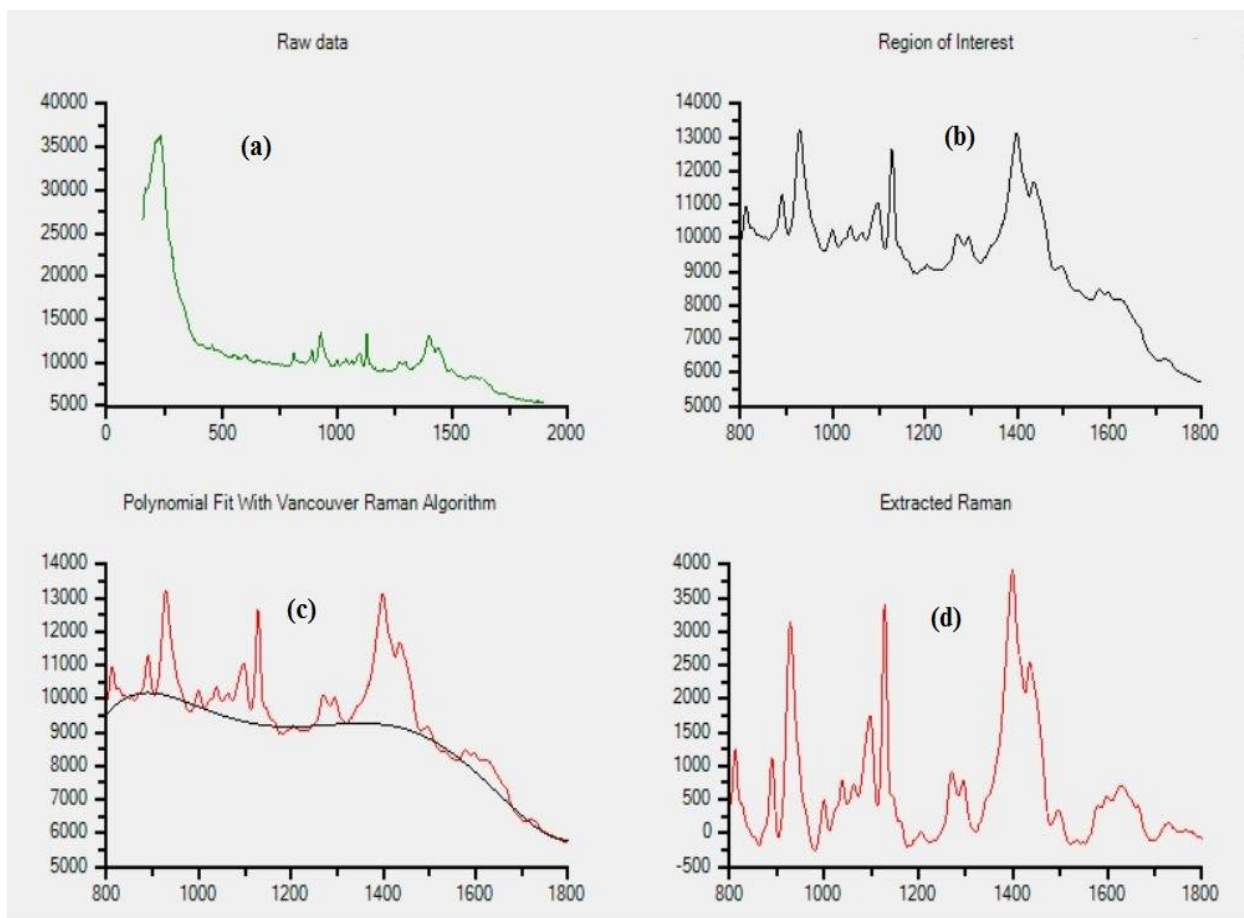
Figure 9: Spectral Profile of HIV-1 p24 antigen + preservative RPMI solvent adsorbed on Silver substrate; Raw Raman spectral data (a), Selected Region of Interest from Raw data (b), Base line Polynomial fit to be subtracted - black line (c) and extracted Raman Signal (d)

To obtain the characteristic Raman spectrum of HIV1-p24 antigen alone, the spectra of RPMI was subtracted from the spectrum of HIV1-p24 + RPMI as displayed in Figure 10 a. With the characteristic spectra of HIV-1 p24 antigen known, the potential of Raman spectroscopy in early HIV-1 detection was further demonstrated by intentionally contaminating healthy plasma with p24 antigen after which spectra of p24 antigen was also obtained by subtracting combined spectra of plasma + RPMI from that of plasma + RPMI + HIV-1 p24 antigen as shown in figure 10 b. Despite the diminished intensity and about 10 cm^{-1} band shift, the spectral profile in figure 10 a and b were identical indicating that the Raman scattered radiations in either of the case emanated from the same source believed to be p24 antigen. The small energy shift and the differences in the intensities can be attributed to the influence of other surrounding molecules which affect the intermolecular bond and hence the vibrational energies in each case.
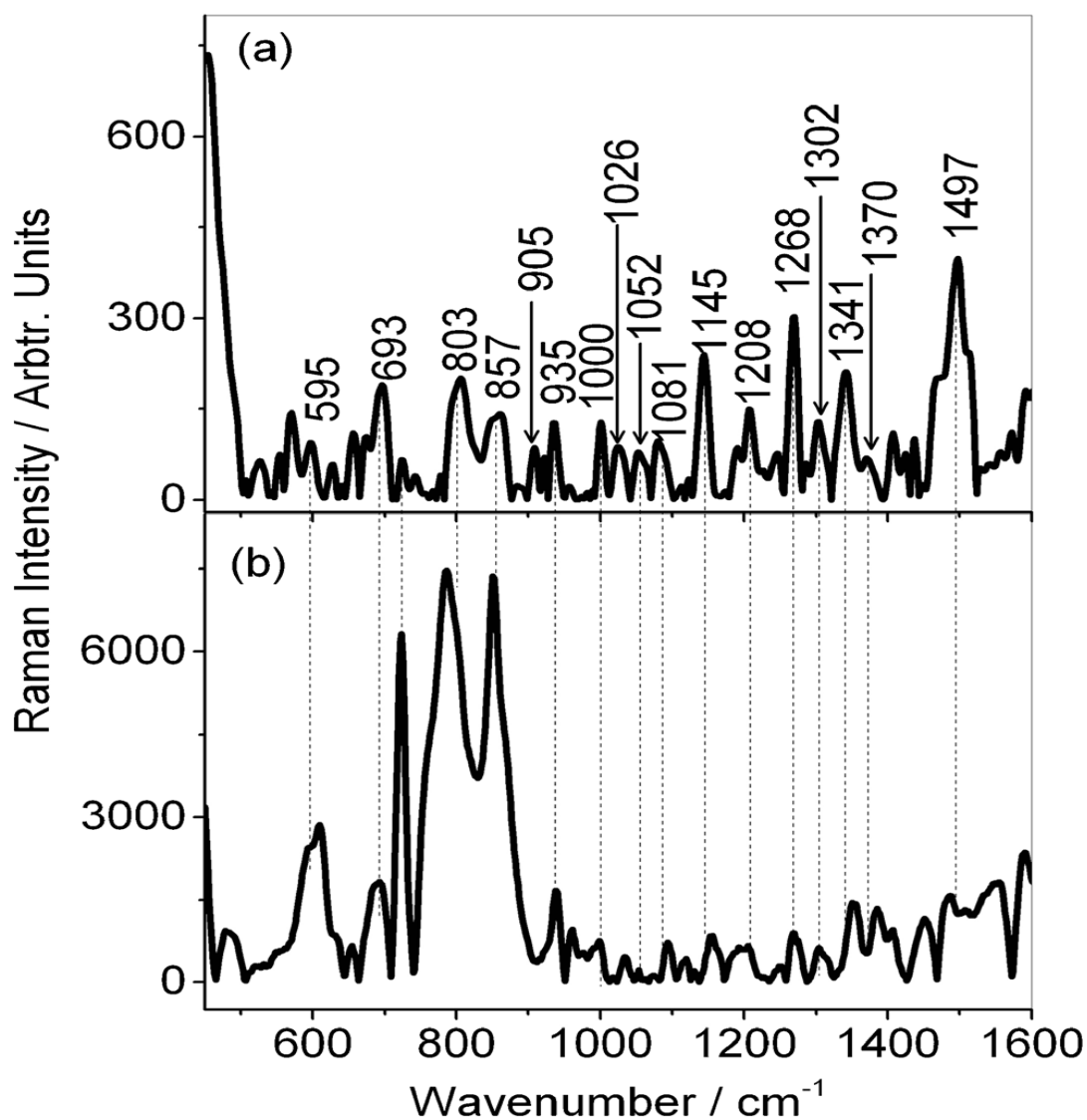
Figure 10: Raman Spectral profile of HIV-1 p24 antigen after 785 nm laser excitation. (a) Shows absolute spectrum after subtraction of spectrum of RPMI from the combined spectra of HIV-1 p24 antigen + RPMI and (b) shows p24 antigen spectrum obtained by subtracting spectra of plasma + RPMI from that of combined plasma + RPMI + HIV-1 p24 antigen (Otange *et al.*, 2017)

The assignment of the peaks to the corresponding vibration and bonds were done in reference to similar studies reported by Lee-Ho and co-workers, (2015) and Pavel *et al*,.(2011), as shown in Table 2. The bands attributed to the RNA were centered at wavenumbers; 595, 1052, 1302 and 1341 cm^{-1} while proteins associated peaks dominated majority of the bands as shown in table 2.

Table 2: Tentative p24 antigen Raman profile peak assignment of prominent bands

Peak position (cm⁻¹) (Based on this work)	Proteins	RNA	Reference
595		Cytosine	(Chuanzong &Yiming, 2005)
693			
803			
857			
905	ν (C-C)		(Chuanzong &Yiming, 2005)
935	ν (C-C)		(Chuanzong &Yiming, 2005)
1000	vs (C-C) Phenylalanine		(Pavel *et al*,. 2011)
1026			
1052		Ribose	(Chuanzong &Yiming, 2005)
1081	ν (C–N) and ν (C–C)		(Pavel *et al*,. 2011)
1170	Tyrosine bending (C-H)		(Pavel *et al*,. 2011)
1208			
1268	Amide III		(Pavel *et al*,. 2011)
1302	ν (C=H)	Adenine, Cytosine	(Chuanzong &Yiming, 2005)
1341	Tryptophan	Adenine	(Pavel *et al*,. 2011)
1370	Tyrosine		(Pavel *et al*,. 2011)
1449	ν (C=H)		(Pavel *et al*,. 2011)
1497			
ν –Stretching Vibration		vs - symmetric stretch	

4.1.2 Principal Component Analysis of the HIV-1 p24 contaminated and uncontaminated plasma Raman spectral data

The great potential of Raman spectroscopy for use in HIV-1 screening within the window period (<14 days after exposure) was demonstrated when used with PCA on the Raman dataset obtained from human plasma intentionally contaminated with HIV1-P24 antigen and from uncontaminated plasma (see Figure 11). PCA allows latent spectral differences to be used in

segregation. PCA uses new axes (which are combinations of height and weight) called principal components (PCs) to group data with similar spectroscopic patterns together using one value called a score per dataset. We found that first two PCs (PC 1and PC 2) accounted for 96.69% of the total variance of the original matrix. The intention was to segregate between the spectra of plasma with p24 antigen (representing plasma recently exposed) and healthy plasma thus demonstrating the applicability of the method in early detection.
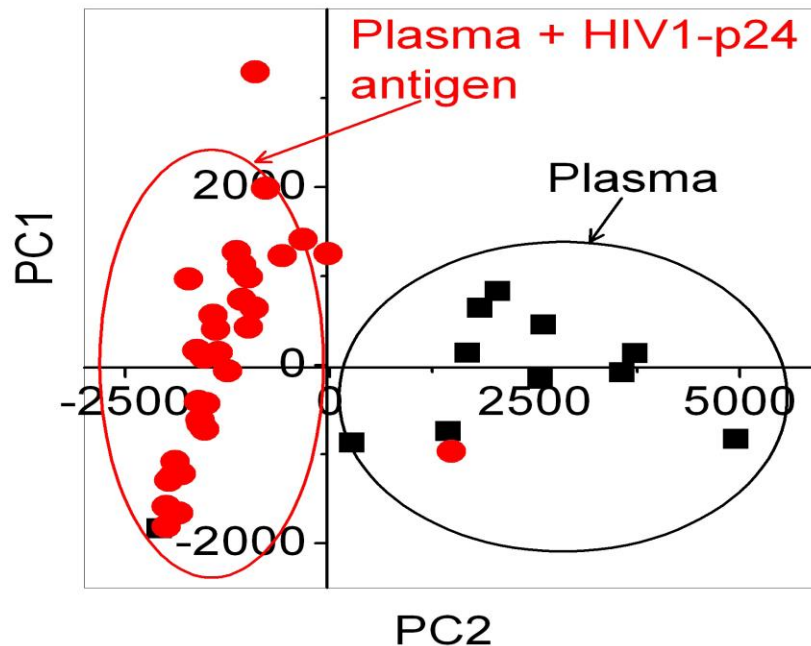
Figure 11: Plots of the first principal component (PC 1) versus the second principal component (PC 2) for spectral data obtained from negative plasma and plasma contaminated with p24 antigen (Otange *et al.*, 2016).

As can be seen in Figure 11, plasma samples contaminated with HIV1-p24 antigen could be clearly distinguished from those without the antigen with sensitivity of 96.5% (28/29) and specificity of 91% (10/11). This indicated that the two Raman spectral data sets had differing spectral patterns thus demonstrating great power of Raman spectroscopy together with PCA in early screening for HIV-1. The samples containing HIV1-p24 antigen had mainly negative PC2 scores while those without the antigen had positive PC 2 scores in the score plot.

4.2 Raman Spectroscopic Characterization of HIV- and HIV+ human blood Plasma

Averaged spectra of HIV- plasma, HIV+ plasma (at selected viral loads; red line - 193, blue – 2753, green – 21771, purple – 37368 and grey – 835020 copies per ml of plasma) and

silver paint smeared glass substrate within the spectral finger print region 350-1750 cm^{-1} are displayed in figure 12 and their tentative peak assignment in table 3. The peaks in the spectra are associated with lipids (713, 928, 1446 cm^{-1}), proteins (645, 713, 813, 1270, 1446, 1658 cm^{-1}) and carbohydrates (928, 990 cm^{-1}) (Virkler and Lednev, 2009; Feng *et al.,* 2010; Feng *et al.,* 2011; Lin *et al.,* 2012 and Feng *et al,.* 2013). Most Raman bands in HIV+ plasma were similar to those in HIV-, however, HIV+ samples displayed unique peaks centered at wavenumbers 1270 and 1446 cm^{-1}. These peaks were assigned to Raman active vibrations in the HIV virion proteins (Virkler and Lednev, 2009; Pavel *et al.,* 2011 and Lee-Ho *et al.,* 2015). It is known that most part of the virion is composed of proteins (60-70 %) followed by lipids (30-40 %), carbohydrates (2-4 %) and nucleic acids (1-2.5%) in that order (Chuanzong and Yiming, 2005).
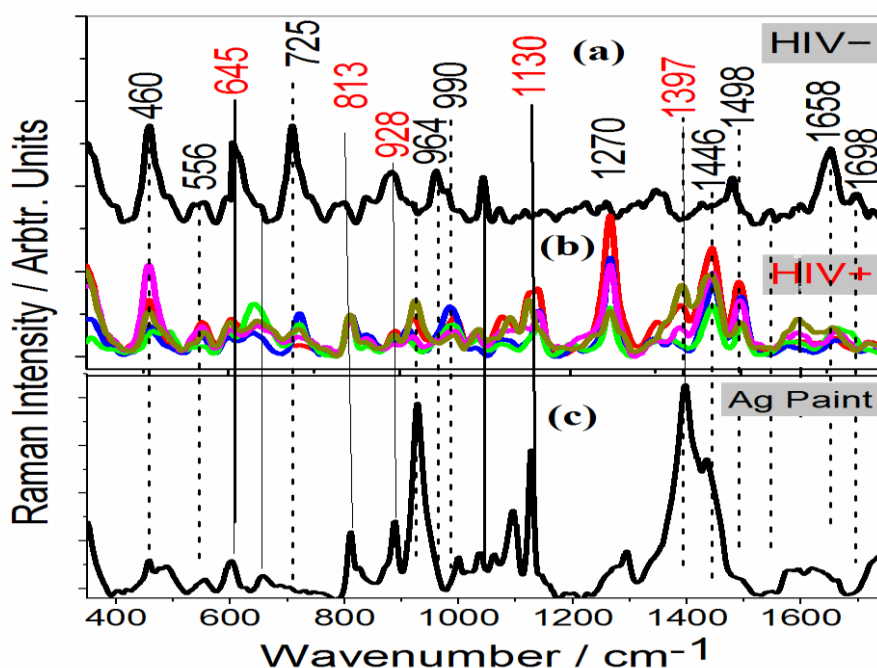
Figure 12: Raman Spectral Profile of HIV- Plasma (a), HIV+ plasma (b) after baseline correction and Silver substrate (c) – inset. HIV- spectra was vertically shifted for clarity (Otange *et al*., 2017).

The proteins available in plasma of both HIV- and HIV+ includes serum, globulins, albumins and fibrinogen, thus the Raman bands associated with them are expected in both their spectra. Signal enhancement was believed to have been caused by the attachment of the samples on silver nanoparticles substrate. Though sparse (35-65 % per weight in the paste), the presence of silver resulted in the strong Raman signals of associated proteins. According to Jose *et al*.,

(2005), the HIV virus has a preferential attachment with the silver nanoparticles through gp120 glycoprotein knobs. The two intense peaks centered at wavenumbers 1270 cm[-1] and at 1446 cm[-1] were ascribed to vibrational state of amide III of α-helix and CH_2 bending vibrational mode in proteins and lipids components of plasma respectively (Feng *et al.,* 2010; Feng *et al.,* 2011; Lin *et al.,* 2012 and Feng *et al.,* 2013). These results were similar to those reported by Chuanzong and Yiming, (2005) from serum samples (plasma without clotting factor) showing that using the cheaper silver paste smeared glass substrates works equally good.

Table 3: SERS peak positions and vibrational mode assignments (Feng *et al.,* 2010; Feng *et al.,* 2011; Lin *et al.,* 2012 and Feng S *et al.,* 2013).

HIV- (cm[-1])	HIV+ (cm[-1])	Band Assignment	Vibrational mode
460	460	unknown	unknown
556	556	Unknown	Unknown
638	638	v(C-S)	L-tyrosine (Protein)
713	missing	v(C-S) and vs (C-N)	Proteins ; Lipids
725	725	unknown	Adenine (RNA)
813	813	unknown	RNA; Alanine
928	928	δ(COH) and v(C-C)	Carbohydrates; Lipids
964	missing	unknown	unknown
990	990	CH_2 Rocking	Carbonydrates
1206	1206	Ring Vibration	Tyrosine, phenylalanine
missing	1270	δ(CH_2)	Amide III (Proteins)
missing	1446	v(CH_2), δ (CH_2) and v(CH)	Proteins and Lipids
1498	1498	vs (N H_3)	Spermine Phosphate hexahydrate
1658	missing	v(C=O)	Amide I
1698	missing	Unknown	Unknown
v -Stretching Vibration		δ – Bending Vibration	vs - symmetric stretch

The other significant difference in the spectra of negative and positive plasma samples occur in the peak intensity at 1658 cm[-1] corresponding to the stretching vibrational mode of C=O bond in amide I of α-helix. This peak is so strong in HIV- plasma and relatively not prominent in

HIV+ plasma (see Figure 12). The spectral profile of HIV+ plasma samples also showed major difference in Raman intensities between the HIV- and HIV+ plasma samples. The degree of this difference in intensity varied depending on the level of viral load with spectra corresponding to more viral load having much lower Raman intensity as shown in Figure 13.
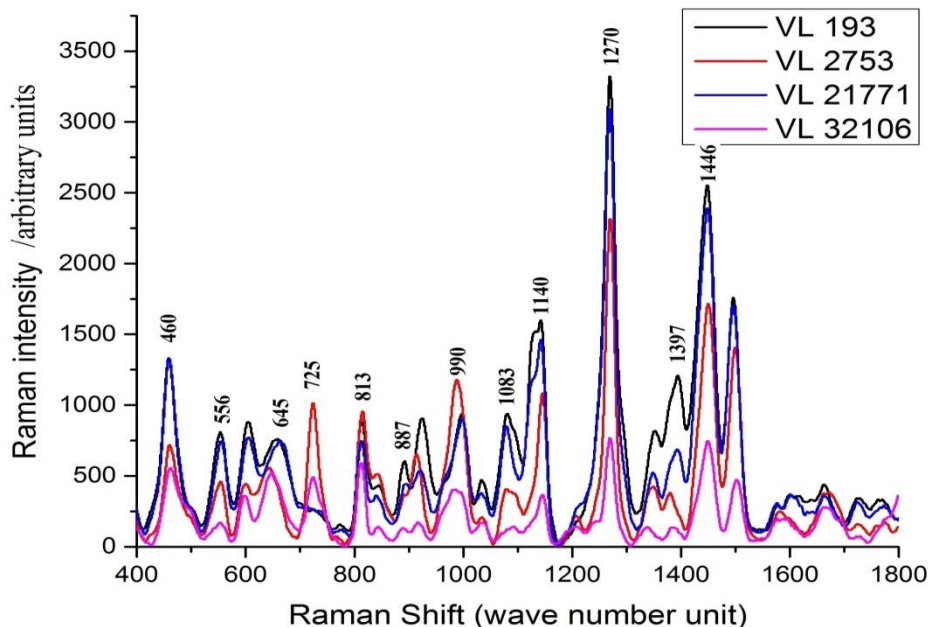
Figure 13: Spectral profile of positive plasma samples of different viral load (number/ml) showing how Raman intensity of HIV+ spectra varies with level of infection.

4.3 Raman Spectroscopic Characterization of HIV- and HIV+ human whole blood

The spectral profile for blood showed major differences other than the decrease in Raman intensity for HIV-1 infected blood. The main changes related to the proteins can be observed at 930, 1206, 1275 and 1445 cm^{-1} (Virkler and Lednev, 2009; Feng *et al.,* 2010; Feng *et al.,* 2011; Lin *et al.,* 2012 and Feng *et al,*. 2013). Despite the drop in intensity of HIV+ blood spectra, the peaks at 477, 958, 1127, 1206, 1445 and 1583 cm^{-1} are more prominent than in the corresponding HIV- blood spectra (see Figure 14). Moreover, after baseline correction, the blood spectral bands centered at 808, 1057, 1159, 1275, 1370, 1396 and 1503 cm^{-1} are still prominent in HIV- spectra than HIV+ spectra (see Figure 15).

Proteins and lipids content reduction in HIV-1 whole blood and plasma samples can be attributed to increased breakdown in the absence of adequate energy intake when they act as alternative fuel sources, poor dietary intake, malabsorption, reduced food intake as a result of poor appetite as well as imbalance between anabolism and catabolism (Pencharz and Jeejeebhoy,

1979 and Pencharz *et al,.* 1981). The proportion of blood components including proteins, lipids and or carbohydrates that is lost as a result of HIV-1 infection depends on the underlying nutritional states, dietary intake as well as the level of viral infection (Macallan *et al,.* 1993).
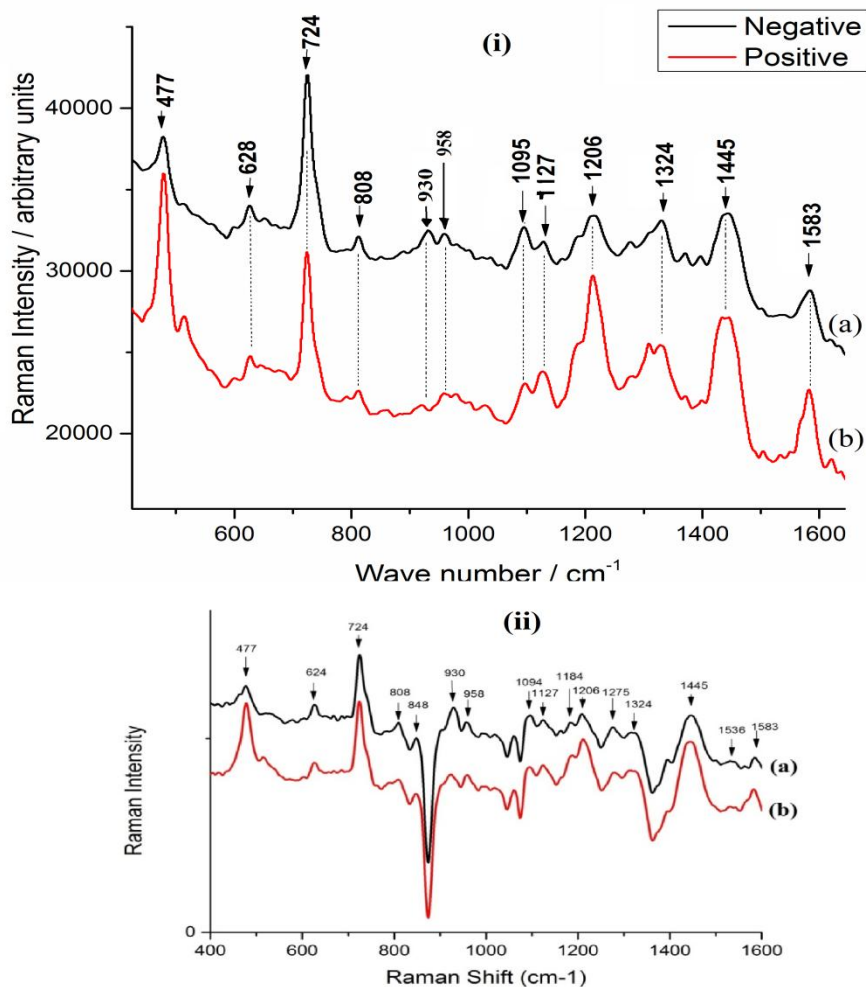
Figure 14: Raman spectral profile of negative blood (a) and positive blood (b) before subtraction of spectral profile of Ag substrate (i) and after spectral subtraction (ii). In both cases, unique profile is observed at wave numbers centered at 477, 930, 958, 1206, 1275 and 1445 cm^{-1}.
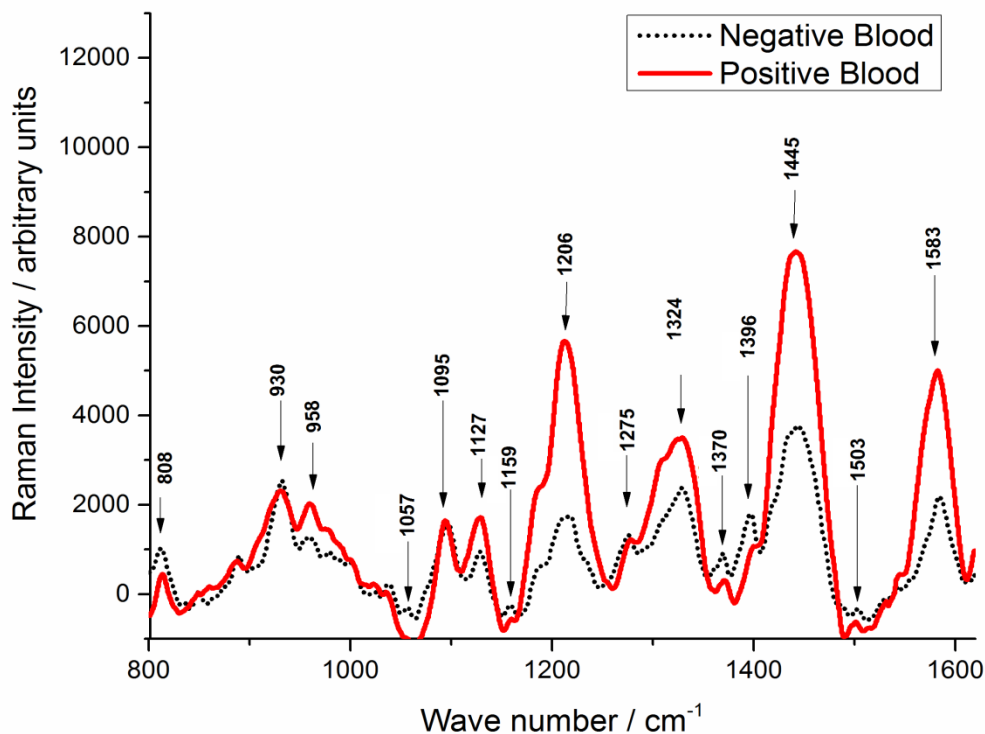
Figure 15: Raman spectral profile of negative and positive blood after baseline fit showing the variation of intensity of peaks based on the same base line.

4.4 Chemometric analysis of Raman spectral data

Other than characterization based on unique spectral peaks, we further sort the application of chemometric techniques; PCA and ANN to segregate between the two groups of spectral dataset (from healthy and infected samples). PCA analyzes a group of spectral dataset and clusters them based on the unique patterns inclusive of latent ones hence facilitating batter segregation and in addition it quickly clusters the spectra once without having to independently verify the unique patterns per spectrum. On the other hand ANN trains the network using known spectral dataset then using the training weights in testing any spectrum dataset from unknown sample.

4.4.1 Segregation of HIV- from HIV+ Raman dataset from plasma samples using Principal Component Analysis

In addition to early detection, Chemometric Raman was also applied for detection of HIV-1 infection at any stage with the application of PCA on the Raman dataset obtained from human plasma infected with HIV-1 and from healthy plasma samples (Figure 16a). PCA

41

algorithm was run on the data matrix after data preprocessing. The variance in the datasets were used in the segregation where linear combination of wavenumbers was formed and ranking them in order of variance. The resulting multivariate dataset is represented by a set of orthogonal axes called principal components (PCs). The first PC (PC1) describes the most variance within the data set followed by PC2 which is orthogonal to the first PC and so on. Each of the spectral data plotted as a point (a score) on a set of two PCs (score plot). Each PC is associated with a spectrum loading vector and its loadings plot has the appearance of a spectrum (Alexander and Tapani, 2010; Herve and Lynne, 2010). We found that first two PCs accounted for 99.76% of the total variance with PC 1 (99.03% variance) and PC 2 (0.73% variance) of the original plasma data matrix as shown in table 4.

Table 4: Eigenvalues and corresponding percentage of variance of PCs

Principal Component	Eigenvalue	Percentage of Variance (%)	Cumulative (%)
1	15.844600	99.03	99.03
2	0.117620	0.73	99.76
3	0.303300	0.19	99.95
4	0.003150	0.02	99.97
5	0.001450	0.01	99.98
6	0.000931	0.01	99.99
7	0.000453	0.01	100.00
Total		**100.00**	**100.00**

Score plots of PC 1 and PC 2 were able to highly discriminate the spectra of plasma from infected group (n=48) and control group (n=28). As can be seen in Figure 16, plasma samples from control samples could be clearly distinguished from those obtained from infected volunteers without using any labeling probe. This was achieved with sensitivity of 100% (48/48) and specificity of 89.28% (25/28) indicating the applicability of Raman spectroscopy together with PCA in a label free detection of HIV-1 infection. Majority of infected samples had mainly negative PC 1 scores while the control samples had mainly positive PC 1 scores in the score plot.

To find out which Raman peaks (bands) were responsible for segregation of the two spectral data sets in PCA, a plot of loadings of each PC versus wavenumbers was done as displayed in Figure 16 (b). The peaks with the largest amplitude in the loadings plot had a greater influence in the differentiation were identified at wavenumbers 460 and 727 cm^{-1} in HIV- and 460, 1002, 1140, 1270 and 1497 cm^{-1} in HIV+ plasma spectra.
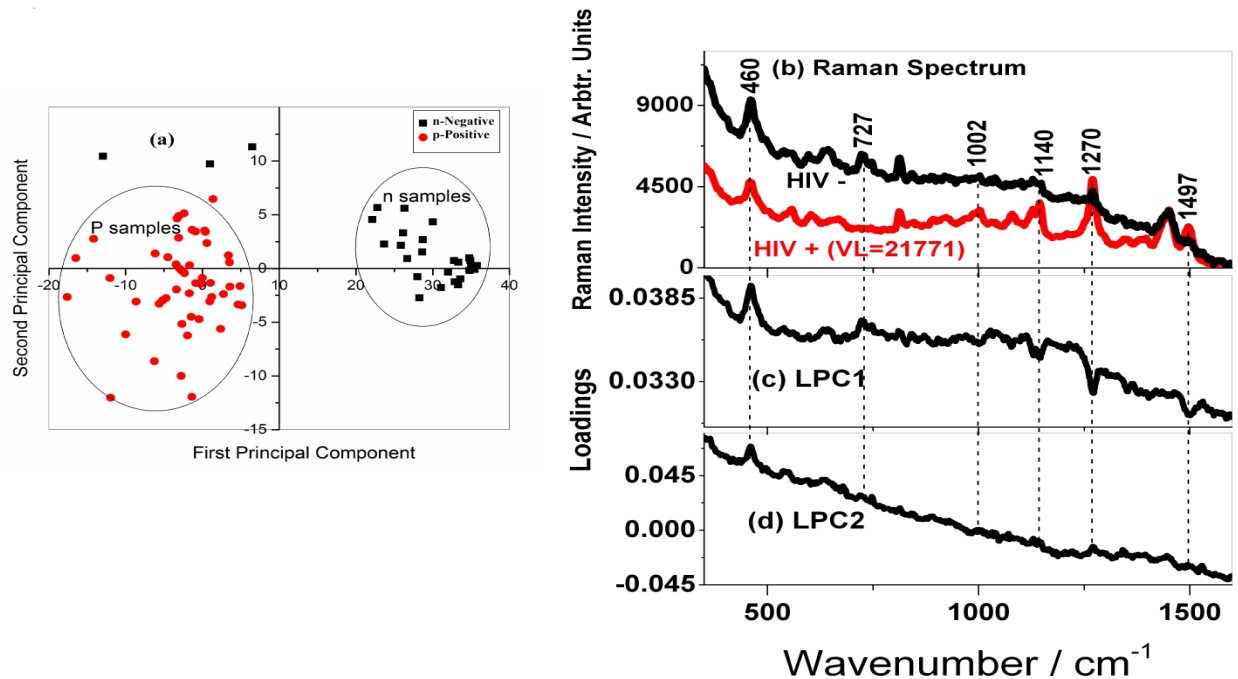
Figure 16: Plots of the first principal component (PC 1) versus the second principal component (PC 2) for healthy group and HIV-1 infected group plasma samples (a), Selected Raman spectrum of HIV- (black line) and HIV+ plasma (red line) (b), Loading plots of PC 1 (c) and Loading plots of PC 2 (d). The equation of the line separating the two groups is PC 1 = 10 (Otange *et al.*, 2017)

4.4.2 Principal Component Analysis on plasma spectral data after auto fluorescence background subtraction

PCA was also applied on Raman dataset from positive and negative plasma after Vancouver Raman Algorithm processing to remove auto fluorescence in the raw data. The first three PCs accounted for 83.2% of the total variance of the Vancouver Raman data matrix (PC 1 accounted for 51.7%, PC 2 accounted for 27.7% and PC 3 accounted for 8.8%). To find out which Raman peaks were responsible for segregation of the two spectral data sets in PCA, a plot

of loadings of each PC versus wavenumbers was done as displayed in figure 17. Figure 17 shows that the peaks with the largest score (whether positive or negative) in the loadings plot had a greater influence in the differentiation are those centered at wavenumbers 460 and 1658 cm⁻¹ in HIV- plasma and 645, 725, 1270, 1498 cm⁻¹ in HIV+ plasma. The assignments of these peaks are given in Table 3.
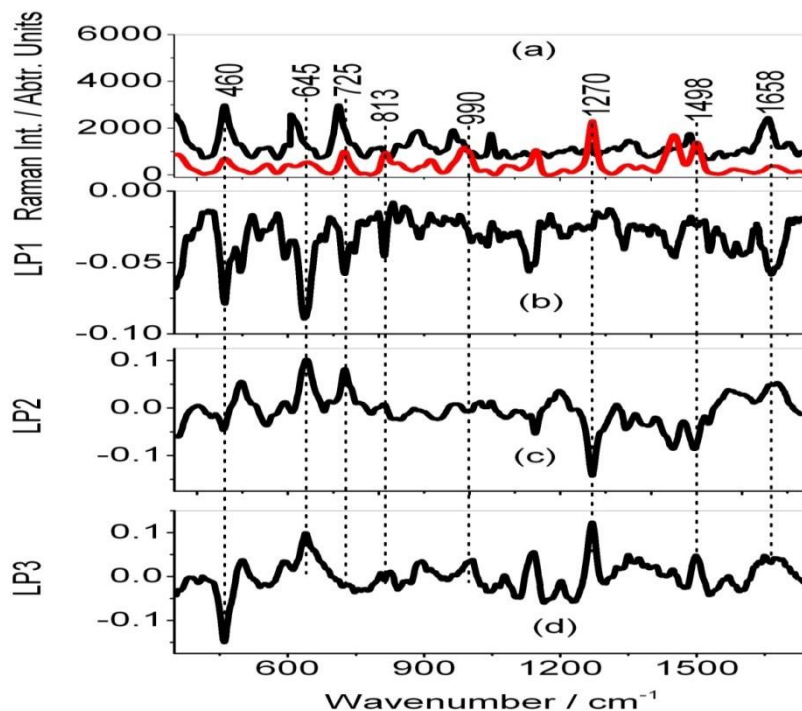
Figure 17: Vancouver Raman spectra of selected HIV negative (black line) and positive (red line) plasma (a), Loading plots of PC 1 (b) PC 2 (c) and PC 3 (d) plotted against the wave numbers. The spectral profile of negative plasma was vertically shifted for clarity (Otange *et al.*, 2016).

The three loading plots were then used to run PCA on Vancouver Raman data matrix for classification of positive and negative plasma. Figure 18 shows that the higher the PC percentage with respect to the total variance of the data matrix, the better the classification of the spectra. This explains why plots of PC 1 against PC 2 (higher percentage of variability) produced better classification as compared to the other plots.
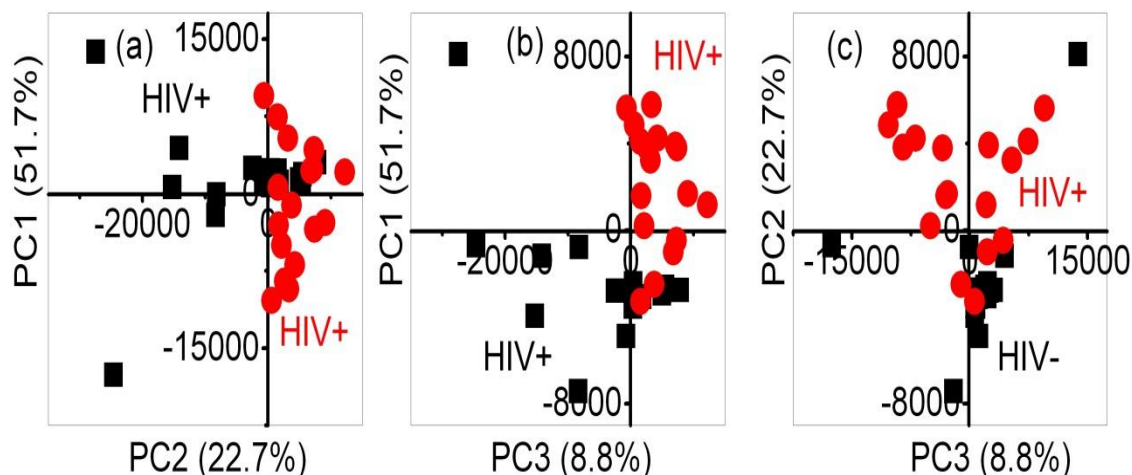
Figure 18: Plots of PC 1 against PC 2 (a), PC 1 against PC 3 (b) and PC 2 against PC 3 (c) for healthy group (black shaded squares) and HIV-1 infected group (red shaded circles) plasma samples (Otange *et al*., 2013).

Score plots of PC 1, PC 2 and PC 2 were able to discriminate the spectra of plasma from infected group and control group. Sensitivity of discrimination decreased with decrease in percentage of variability of the principal components. In comparison to the PCA analysis before background subtraction (section 4.4.1), segregation was more clear when auto fluorescence background data was not subtracted. The removal of background auto fluorescence reduces the degree of the differences between the two groups of spectra. Consequently, this interferes with the percentage of variability of all the principal components especially PC 1 which is greatly reduced. The reduction in PC 1 variability values could explain the poor segregation comparatively seen after spectral subtraction. However, with regard to PC 2 whose percentage of variability increases greatly after removal of auto fluorescence, segregation of spectral groups becomes better. This is observed by comparing the most discriminating PC in figure 16 and figure 18a and c.

4.4.3 Principal Component Analysis of blood samples

The potential of chemometric SERS for detection of HIV-1 infection in blood was also demonstrated with the application of PCA on the Raman data set obtained from human blood infected with HIV-1 (n=18) and from healthy blood samples (n=18). PCA algorithm was run on the data matrix after data preprocessing. We found that first two PCs accounted for 99.89% (PC 1 = 99.64% and PC 2 = 0.25%) of the total variance of the original data matrix (see table 5).

45

Table 5: Eigenvalues and corresponding percentage of variance of PCs

| Principal Component	Eigenvalue	Percentage of Variance (%)	Cumulative (%)
1	48.825220	99.64	99.64
2	0.121190	0.25	99.89
3	0.017710	0.04	99.93
4	0.010330	0.02	99.95
5	0.007670	0.02	99.96
6	0.005690	0.01	99.98
7	0.001490	0.01	99.99
8	0.000232	0.01	100.00
Total		**100.00**	**100.00**
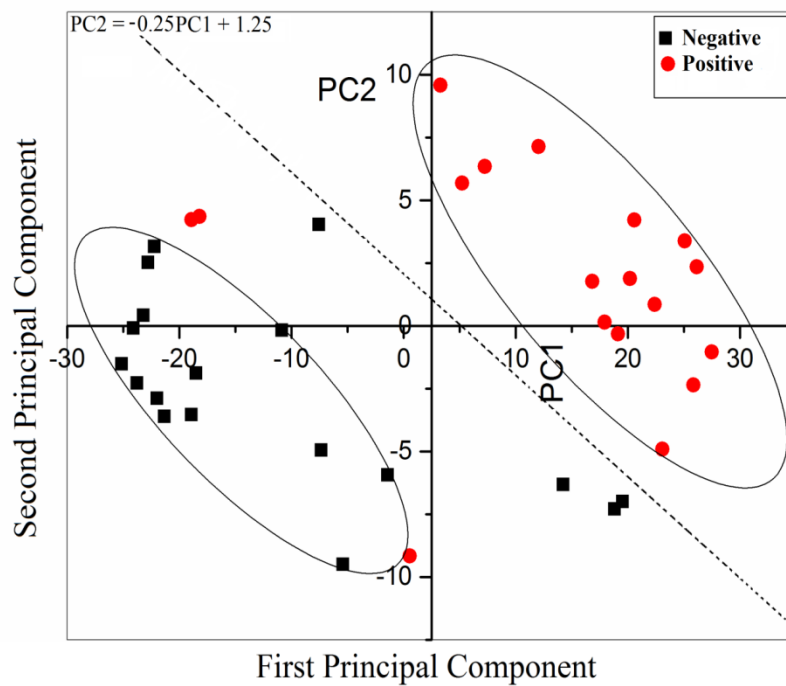
Figure 19: Plots of the first principal component (PC1) versus the second principal component (PC2) for healthy group (black shaded squares) and infected group blood samples (red shaded circles). The equation of the line separating the two groups (dotted line) is PC 2 = -0.25 PC 1+ 1.25.

Score plots of PC 1 and PC 2 was highly discriminating for the spectra of infected group and control group. As can be seen in figure 19, blood samples from negative volunteers could be clearly distinguished from those obtained from infected volunteers without using any labeling probe. This was achieved with sensitivity of 83.3% (15/18) and specificity of 100% (18/18) indicating the applicability of Raman spectroscopy together with PCA in label free detection of HIV-1 infection. Majority of infected samples had mainly positive PC 1 scores while the control samples had mainly negative PC 1 scores in the score plot.

4.4.4 Prediction of HIV-1 infection Using Artificial Neural Network

An ANN technique was also evaluated to predict HIV-1 infection from SERS dataset of both infected and control whole blood as well as their corresponding plasma. When presented with the testing set, the network generated an output with correlation coefficient value in the range of 0.0 (0%) $\leq R \leq 1.0$ (100%) reflecting its predictive value for HIV-1 infection. The square of correlation coefficient R^2 can be defined as the ration of explained variation to the total variation of the matrix dataset (Floyd and Tourassi, 1992). 0% indicates that the model explains none of the variability of the response data around its mean while 100% indicates that the model explains all the variability of the response data around its mean. The ANN training and testing model resulted in $R^2 = 0.99994$ for HIV+ plasma sample and $R^2 = 0.9601$ for HIV+ whole blood samples indicating that the detection achieved clinically relevant precision (see Figure 20 and Figure 21). Generally, R^2 values higher than 0.9 indicate that the method under investigation is clinically accurate (Annika *et al*., 2002). Based on this criterion, our method of HIV detection was clinically accurate. However, when a network trained using infected Raman dataset was used to test a control Raman dataset of the same fluid type as a target, the value of square of correlation coefficient from the trained network was approximately equal to zero. Similarly, when a network trained using control Raman dataset was used to test an infected Raman dataset of the same fluid type as a target, the value of correlation coefficient from the trained network resulted in $R^2 \approx 0$ (see figure 22)
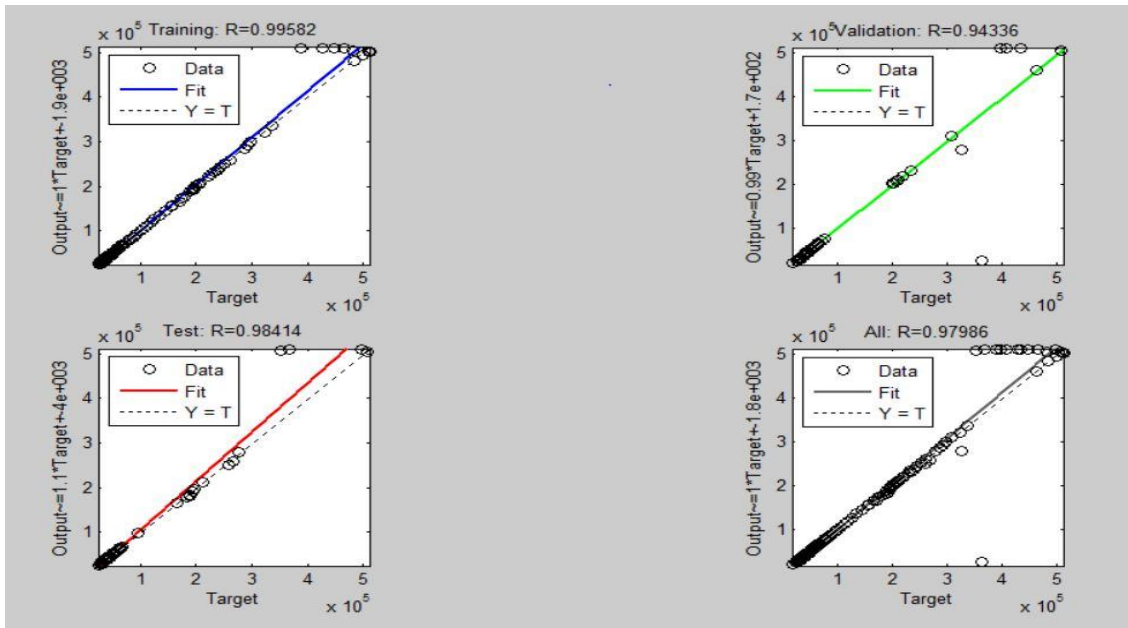
Figure 20: ANN correlation output from HIV+ whole blood Raman dataset. Training and testing Raman dataset from HIV+ whole blood achieved a diagnostic regression $R^2 = 0.9601$ (R = 0.97986)
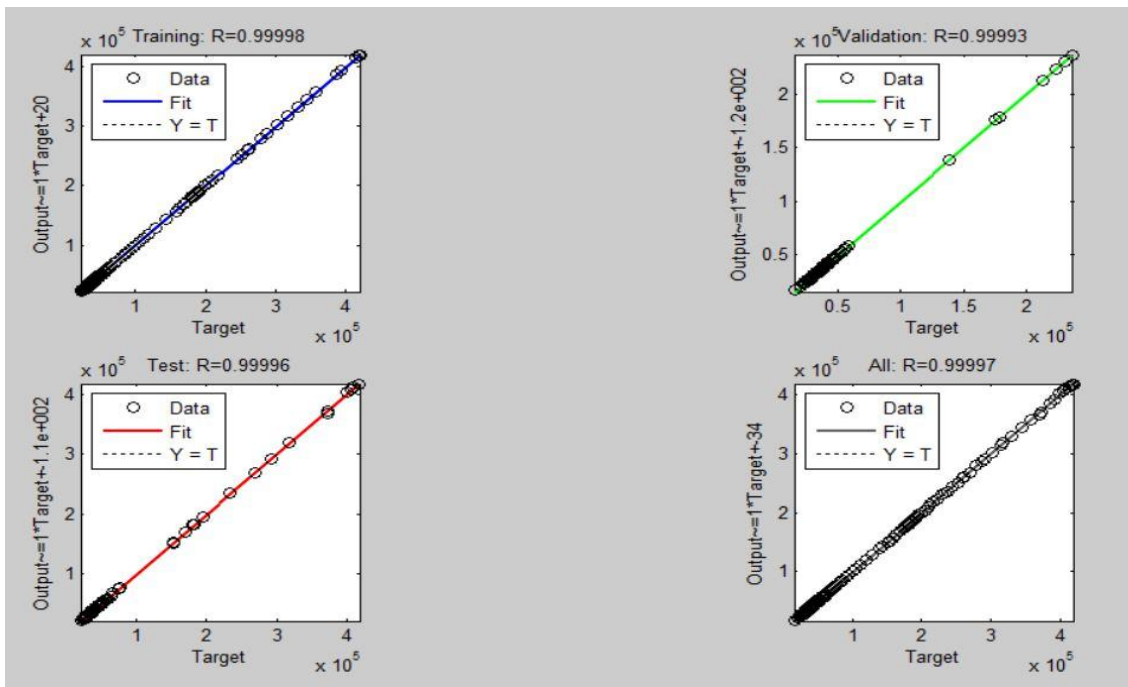
Figure 21: ANN correlation output from HIV+ plasma Raman dataset. Training and testing Raman dataset from HIV+ plasma achieved a diagnostic regression $R^2 = 0.99994$ (R = 0.99997)
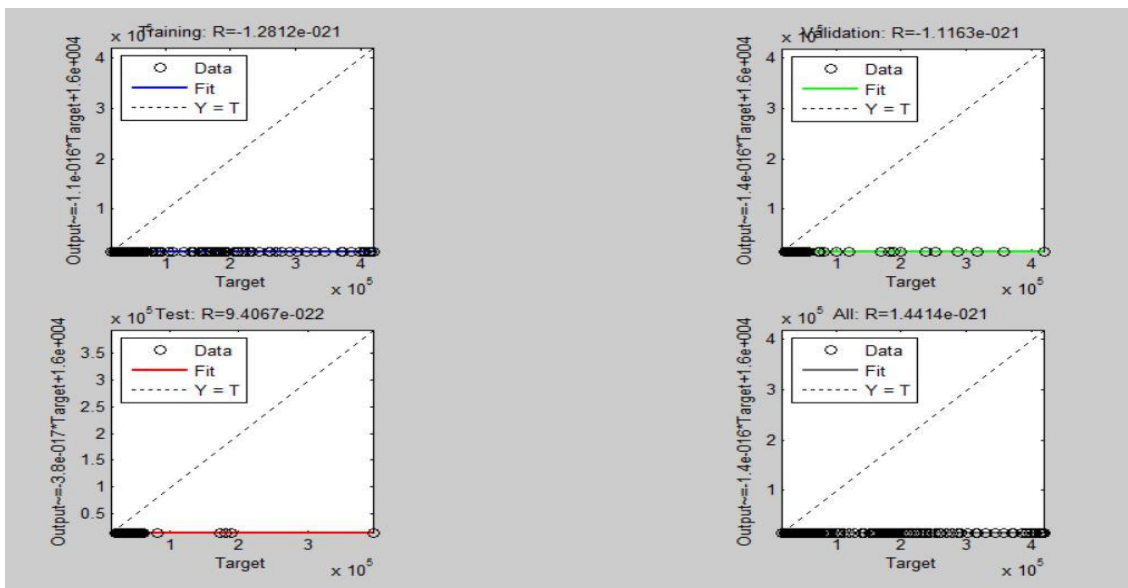
Figure 22: ANN correlation output Trained using infected plasma Raman dataset but tested on control plasma Raman dataset.

4.5 Quantitative Analysis of HIV-1 Viral Load based on Prominent Peaks of HIV+ plasma spectra

The dominant Raman peaks observed and believed to exist as a result of HIV-1 infection in plasma were 1270 cm^{-1} and 1446 cm^{-1}. These peaks were similar to those reported by Chuanzong and Yiming, (2005) from serum samples and were ascribed to vibrational state of amide III of α-helix (proteins) and CH$_2$ bending vibrational mode in proteins and lipids components of plasma respectively (Feng *et al.,* 2010; Feng *et al.,* 2011; Lin *et al.,* 2012 and Feng *et al.,* 2013). These peaks (1270 cm^{-1} and 1446 cm^{-1}) were examined to determine the relationship between the HIV-1 concentration and peak intensity. The decrease in Raman intensity was presumably caused by the effect of HIV-1 infection on the plasma components. The Raman intensity of each of these unique peaks was used to develop a linear estimation correlated with the viral load concentration. The linear relationships between the inverse of logarithmic of the component peak against the viral load concentration are shown in Figure 23 for both peaks as well as the magnified peaks centered at 1270 cm^{-1} and 1446 cm^{-1}.

In examining the relationship between the inverse logarithmic of component peaks and viral load concentration, the calculated viral load of the leave out cross validation showed slight variation believed to be caused by none uniformity of the Ag nanostructure. After infection, HIV is believed to spread within the body exponentially. The effect of viral load on the level of body

49

components is inversely proportional, thus the higher the viral load, the lower the Raman peak intensity for infected samples. In order to come up with a linear model that relates peak values directly to viral load, we calculated the inverse of logarithmic value of Raman HIV-1 associated peak intensity and plotted it against viral loads from the following relation.

$$y^{-1} \alpha \, \exp(x)$$

Where y is Raman peak intensity and x is HIV viral load. Introducing natural logarithm on both sides of the above equation gives

$$lny^{-1} \alpha \, x$$

But from the definition $lny^{-1} = \dfrac{\log y^{-1}}{\log e}$

Therefore, $\log y^{-1} \alpha \, xloge$

Since $loge$ is a constant, then

$$\log y \, \alpha \, -x$$

The negative sign just before x shows that the two terms are inversely proportional. Thus, to obtain a directly proportional relation between the two terms, the above equation was reduced to: $\left(\log y\right)^{-1} \alpha \, x$

The Raman peak observed at 1270 cm^{-1} (amide III of α-helix) showed a linear relationship to viral load with square of coefficient of correlation $R^2 = 0.9457$. The peak ascribed to the CH$_2$ bending vibrational mode in proteins and lipids at 1446 cm^{-1} also showed a linear correlation to vir3al load with square of coefficient of correlation $R^2 = 0.9959$. These results demonstrate the successful detection (based on the presence of these peaks) and quantification of HIV-1 infection using Raman spectrum. From the criterion given by Annika and his co-workers (2002) on accuracy of clinical procedures, this technique is sufficiently accurate, sensitive and selective to study HIV-1 and its associated effects.
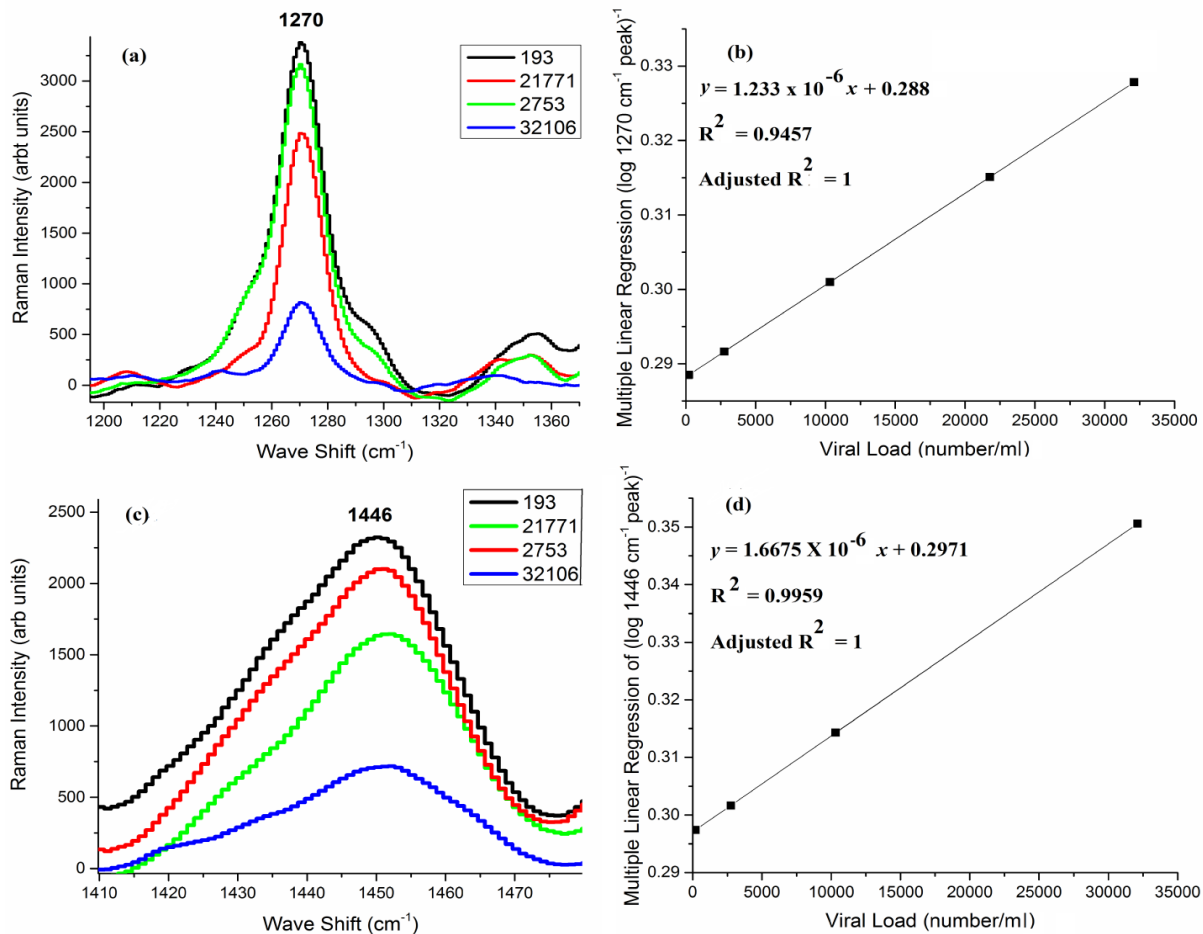
Figure 23: Quantification of HIV-1 viral load based on Raman Intensity of peak centered at 1270 cm^{-1} and 1446 cm^{-1}. Peak centered at 1270 cm^{-1} of selected viral load showing how the Raman spectral intensity varies with viral load (a), Multiple Linear Regression model for quantifying HIV-1viral load in blood plasma based on peak variation centered at 1270 cm^{-1} (b), Peak centered at 1446 cm^{-1} of selected viral load showing how the Raman spectral intensity varies with viral load (c) and Multiple Linear Regression model for quantifying HIV-1viral load in blood plasma based on peak variation centered at 1446 cm^{-1} (d)

4.6 Quantitative Analysis of HIV-1 Viral Load based on Artificial Neural Network Regression

Quantitative estimation of the viral load was also done and the findings displayed in regression plot shown in figure 24. The virus load values predicted by the model versus those obtained from PCR using HIV+ spectral data values around the band centered at 1270 cm^{-1}. High values of R and R^2: 0.9333958 and 0.9895 respectively, were obtained indicating great

potentiality of Raman spectroscopy in estimating the viral load values when ANN is first trained using spectral data around 1270 cm^{-1} band.
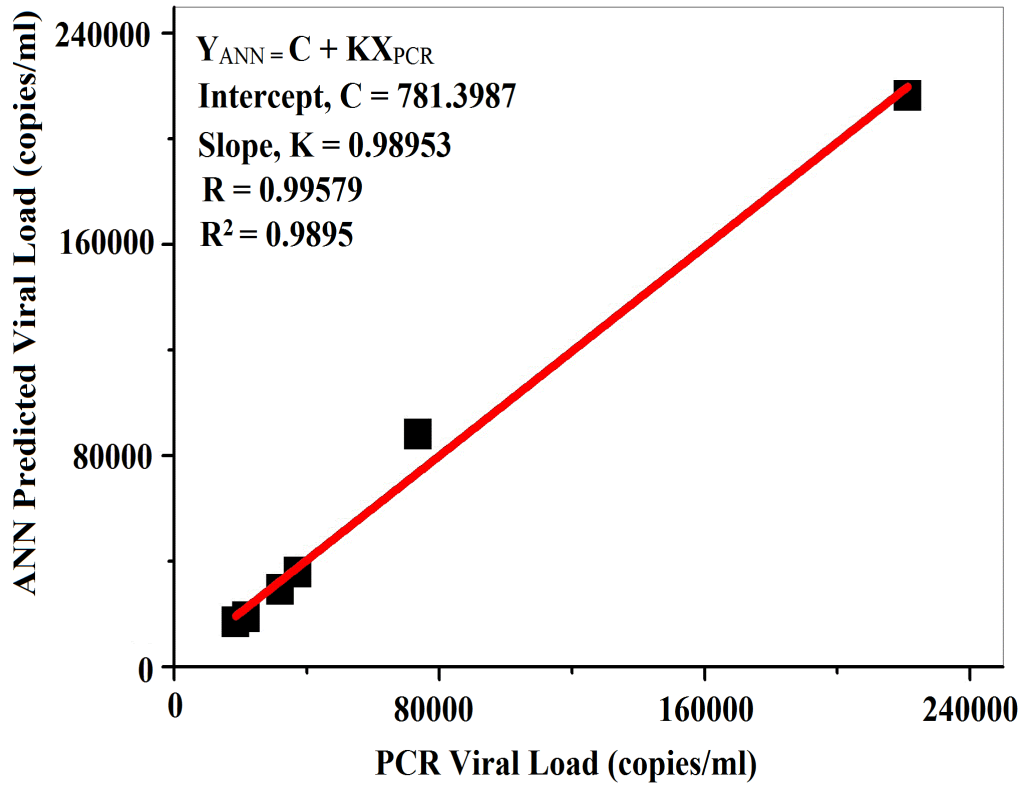
Figure 24: The regression plot showing the relationship between the HIV viral loads value predicted using ANN and those measured using PCR. There was a good correlation between the two results with a R and R^2 values of 0.9958 and 0.9895 respectively (Otange *et al.*, 2017).

CHAPTER FIVE
CONCLUSIONS AND RECCOMMENDATIONS

5.1 Conclusions

5.1.1 Conclusions on Characterizations

A spectroscopic signature unique to the samples were developed using Raman spectroscopy based on the heterogeneous chemical nature of samples. The Raman spectral profiles of p24 antigen, infected samples and control samples could be used to determine the effect of HIV-1 infection on blood and plasma components. Raman spectral profile showed specific biochemical changes and reduction in body components especially proteins and lipids associated with HIV-1 infection in blood and plasma samples. These biomolecular changes were then utilized in classifying the different groups of spectra in relation to infection status. Similarly, the peak profile of p24 antigen could also be used to check possible peak overlap within the infected plasma samples. Therefore, Raman spectroscopy may be a suitable candidate for evaluating HIV-1 infection related changes in blood and plasma samples thus providing useful information that can really help at individual level HIV diagnosis.

5.1.2 Conclusions on Classification

Use of silver paste smeared glass slides as chemometric based Raman spectroscopy sample substrates can provide a label free detection and differentiation between HIV- and HIV+ blood and plasma samples. Statistical analysis suggested that the spectral profile of blood and plasma samples dataset can be classified as either infected or healthy samples based on ANN training and pattern matching algorithm as well as PCA loading plots. The PCA was able to perform this segregation by generating PCs comprising a reduced number of orthogonal variables. Each loading vector related to the original spectrum representing the weight of that particular component against the basis spectrum and is capable of reflecting the differences between different groups. Similarly, ANN training also resulted in distinct Raman dataset from healthy and HIV-1 infected samples based on R^2 value of the input, target as well as the validation dataset. These results showed that Raman spectra from HIV- and HIV+ blood and plasma samples had different spectral patterns. The technique therefore has the potential of being applied in rapid screening of blood samples in sensitive and label free diagnosis of HIV-1 infection.

5.1.3 Conclusions on Quantification

The result showed that the model could successfully estimate the concentration of HIV-1 viral load within the infected plasma samples. The linear quantification model generated provided a clinically accurate estimation of HIV concentration that could be applied to infected plasma samples with unknown viral load and generate results comparable to the reference PCR. Unlike the reference method, the methodology demonstrated in this work is rapid, label free and relatively simple to execute.

5.2 Recommendations

5.2.1 Recommendations for Characterization

Raman spectral profile of biological samples is affected by the type of solvent, fluorescence, signal to noise ratio, temperature and chemical reaction (resulting in either folding/unfolding, oxidation and or reduction). We recommend minimization of fluorescence effects by selection of suitable wavelength, fluorescence base line correction and or photo bleaching of colored samples. For reproducible measurements we also recommend longer signal acquisition and several spectral accumulations to improve signal to noise ratio, measurements to be made at moderate temperatures as well as noticing and taking care of any sample and substrate preparation that may initiate chemical reaction.

5.2.2 Recommendations for Classification

In order to improve the sensitivity and specificity of PCA discrimination between the spectra in classification and regression coefficient of ANN training algorithm, we recommend the future study to attempt using other methods of making SERS substrate for better enhancement and to better reveal the different between healthy and infected spectral samples. We also recommend that the method used in making the SERS substrate should result in uniform nanoparticle network with same size, shape and inter particle spacing between the colloids to ensure a uniform LSPR hence uniform enhancement of SERS signals.

5.2.3 Recommendations for Quantification

Following the observed miss match in viral load and Raman spectral profile intensity, we recommend that the study participants be categorized into similar socioeconomic background. This is because the level of viral load is affecting the intensity of the spectral pattern as a result of the associated effect on the biochemical composition of blood and plasma. It is expected that

the higher the viral load, the lower the intensity of spectral profile. However, this was not observed in all the spectral profile corresponding to different viral load. The main reason behind this observation was the fact that study subjects were not from similar economic background. Subjects from better economic class who had higher viral load but could afford a good balanced diet still had relatively higher than expected intensity of spectral pattern. Similarly, subjects from poor economic background who had lower viral load but could not enrich their diet, will also have relatively lower than expected spectral profile. Apart from similar economic class, we also recommend that the study subjects be under similar treatment for at least one week. Even though, all the study subjects were subjected to an overnight fasting, blood and plasma biochemical contents were still dependent on resent diet and were not only affected by viral load.

6.0 REFERENCES

Ajith Abraham. (2005). Introduction to Artificial Neural Networks. In A. Abraham, & P. H. Thorn (Ed.), *Artificial Neural Networks*. **2**. 901-908. John Wiley & Sons.

Akikazu Sakudo, , Roumiana Tsenkova, Taisuke Onozuka, Kazuo Morita, Shuming Li, Jiranan Warachit, Yukie Iwabu, Guimei Li, Takashi Onodera, and Kazuyoshi Ikuta. (2005). A Novel Diagonistic Method for Human Immunodeficiency Virus Type-1 in Plasma by Near Infrared Spectroscopy. *Microbiol. Immunol, 49*(7), 695–701.

Alexander Ilin and Tapani Raiko. (2010). Practical Approaches to Principal Component Analysis in the Presence of Missing Values. *Journal of Machine Learning Research, 11*, 1957-2000.

Annika Enejder, Tae-Woong Koo, Jeankun Oh, Martin Hunter, Slobodan Sasic, and Michael Feld. (2002). Blood analysis by Raman spectroscopy. *OPTICS LETTERS, 27*(22), 2004-2006.

Anupam Misraa, Lori Kamemotob, Ningjie Huc, Ava Dykesa, Qigui Yuc, Pavel Zinina. (2011). Micro-Raman Spectroscopy study of ALVAC virus infected chicken embryo cells. In E. M. Brian Cullum (Ed.), *Prococeedings of SPIE. 8025-80250C-1*, pp. 1-32. Smart Biomedical and Physiological Sensor Technology.

Avert Group. (2015, 8 24). *History of HIV and AIDS*. Retrieved from Avert: http://www.avert.org/history-hiv-and-aids.htm

Banwell, McCash, Collin N.and Elaine. (2007). *Fundamentals of Molecular Spectroscopy.* New Delhi: Tata McGraw-Hill.

Block Olivia, Anirban Mitra, Lukas Novotny and Carrie Dykes. (2012). A rapid label-free method for quantitation of human immunodeficiency virus type-1 particles by nanospectroscopy. *Journal of virological methods , 182*(1), 70-75.

Busch MP and Satten GA. (1997). Time course of viremia and antibody seroconversion following human immunodeficiency virus exposure. *American Journal of Medicine, 102*(5B), 117-124.

Carlson and Hunterian. (2004). Insulin resistance in human sepsis: implications for the nutritional and metabolic care of the critically ill surgical patients. *Ann. R. Coll. Surg. 86*(2), 75-81.

Chan J, Taylor D, Zwerdling T, Lane S, Ihara K, and Huser T . (2006). Micro-Raman spectroscopy detects individual neoplastic and normal hematopoietic cells. *Journal of Biophys, 90*(2), 648–656.

Chong-Shi. (2002). Simple Preparation Method for Silver SERS Substrate by Reduction of AgNO3 on Copper Foil. *Applied Spectroscopy, 56*(3-305).

Chuanzong, XU Yiming & LU. (2005). Raman spectroscopic study on structure of human immunodeficiency immunodeficiency virus (HIV) and hypericin-induced photosensitive damage of HIV. *Science in China, 48*(2), 117—132.

Coffin Williams. (1999). Molecular Biology of HIV-1. *World Health Organization, 86*(3-40).

Coffin, John, Ashley Haase, Jay A. Levy, Luc Montagnier, Steven Oroszlan, Natalie Teich, Howard Temin, Kumao Toyoshima, Harold Varmus, and Peter Vogt. (1996). What to call the AIDS virus? *Nature, 321*(6065), 10.

David. (2007). Raman Scattering Theory. *Optics, 49*(5-12), 1-13.

David, C. (Producer), & R., J. (Director). (2012). *Raman Spectroscopy for proteins* [Motion Picture]. India: Jobin Yvone. Retrieved from http:www.horiba.com/scientific

Dawson .A, Austin .R, Wienberg .D. (1991). Nuclear grading of breast carcinoma by image analysis: Classification by multivariate and Neural Network analysis. *American Journal of Clinical athology, 95*(1), 29-37.

Dinh Vo, Yan F, and Wabuyele M. . (2005). Surface-enhanced Raman scattering for medical diagnostics and biological imaging. *Journal of Raman Spectrosc, 36*(6-7), 640-647.

Ezzel. (2002). An overview of the challenges of developing HIV vaccine. *Scientific America*(38), 4-16.

Fauci and Anthony. (1993). HIV infection is active and progressive in lymphoid tissue during the clinically latent stage of disease. *Nature, 362*(41), 355-358.

Feng S, Chen R, Lin J, Pan J, Chen G, Li Y, Cheng M, Huang Z, Chen J, and Zeng H. (2010). Nasopharyngeal cancer detection based on blood plasma surface-enhanced Raman spectroscopy and multivariate analysis. *Biosens and Bioelectron., 25*(11), 2414–2419.

Feng S, Chen R, Lin J, Pan J, Wu Y, Li Y, Chen J, and Zeng H. (2011). Gastric cancer detection based on blood plasma surface-enhanced Raman spectroscopy excited by polarized laser light. *Biosens and Bioelectron, 26*(7), 3167–3174 .

Feng S, Lin D, Lin J, Li B, Huang Z, Chen G, Zhang W, Wang L, Pan J, Chen R and Zeng H. (2013). Blood plasma surface-enhanced Raman spectroscopy for non-invasive optical detection of cervical cancer. *Scientific Analyst , 138* (14), 3967–3974 .

Floyd and Tourassi. (1992). An Artificial Neural Network for lesion detection in SPEC images. *Invest Radiology, 27*, 667-721.

Gelder J, Gussem K, Vandenabeele P, and Moens L. (2007). Reference database of Raman spectra of biological molecules. *Journal of Raman Spectroscopy, 38*(9), 1133–1147.

Ghys. (2013). HIV Assays. *World Health Organizations*.

Grant Webstera, Leann Tilley, Samantha Deedb, Don McNaughtona, Bayden Wooda. (2008). Resonance Raman spectroscopy can detect structural changes in haemozoin (malaria pigment) following incubation with chloroquine in infected erythrocytes. *Federation of European Biochemical Societies, 582*(35), 1087–1092.

Greene and Warner. (1993). AIDS and the immune system. *Scientific American*, 99-105.

Grossberg S. (1976). Adaptive Pattern Classification and Universal versal Feature Detectors. *Biological Cybernetics, 23*(7), 121–134.

Gyeong Bok Jung, Kyung-A Kim, Ihn Han, Young-Guk Park and Hun-Kuk Park. (2014). Biochemical characterization of human gingivalcrevicular fluid during orthodontic tooth movement using Raman spectroscopy. *Bio Optics Express, 1364-003508*(5), 1-14.

Hanson. (2015). Package for Chemometric Analysis of Spectroscopic Data. *Chemospec, 7*(351), 1-37.

Herve Abd iand Lynne Williams. (2010). Principal component analysis. *WIREs Computational Statistics, 2*, 433-459.

Jason Guicheteau, Mikella Farrell, Steven Christesen, Augustus Fountain. (2013). Surface-Enhanced Raman Scattering (SERS) Evaluation Protocol for Nanometallic Surfaces. *Applied Spectroscopy, 67*(4), 396-403.

Jensen M. (2011). Global health sector strategy on HIV/AIDS 2011-2015. *World Health Organization, 7*(12), 236-254.

JIANHUA ZHAO, HARVEY LUI, DAVID I. MCLEAN, and HAISHAN ZENG. (2007). Automated Autofluorescence Background Subtraction Algorithm for Biomedical Raman Spectroscopy. *Applied Spectroscopy, 61*(11), 1225-1232.

Jose Luis, Justin Burt, Jose Morones, Alejandra Bragado, Xiaoxia Gao, Humberto Lara and Miguel Yacaman. (2005). Interaction of silver nanoparticles with HIV-1. *Journal of Nanobiotechnology, 3*(6), 1-10.

Juan Hu, Peng-Cheng Zheng, Jian-Hui Jiang, Guo-Li Shen, Ru-Qin Yu and Guo-Kun Liu. (2010). Sub-attomolar HIV-1 DNA detection using surface-enhanced Raman spectroscopy. *The Royal Society of Chemistry, 135*(1), 1085-1089.

Kohonen T. (1984). Self-Organization and associative memory. In J. Kloverd (Ed.), *Series Information Science. 8*, pp. 50-62. Berlin: Springer.

Kotler D, Rosenbaum K, Wang J, Pierson R. (1999). Studies of body composition and fat distribution in HIV-infected and control subjects. *Journal of Acquired Immune.Defic. Syndr. Hum. Retrovirol, 20*(3), 37-228.

Kudelski. (2008). Analytical applications of Raman spectroscopy. *Journal of Raman S[ectroscopy, 76(1)*(1–8).

Kuzmany, H. (2009). *Solid-state spectroscopy: an introduction*. Springer Science & Business Media.

Lee Jin-Ho, Byung-Keun Oh and Jeong-Woo Choi. (2015). Development of a HIV-1 Virus Detection System Based on Nanotechnology. *Sensors, 15*(1424-8220), 9915-9927. doi:**10.3390/s150509915**

Lee Ho, Byung-Keun Oh and Jeong-Woo Choi. (2015). Rapid and Sensitive Determination of HIV-1 Virus Based on Surface Enhanced Raman Spectroscopy. *Journal of Biomedical Nanotechnology, 11*(9), 2223–2230. doi:0.**1166/jbn.2015.2117**

Lin D, Feng S, Pan J, Chen Y, Lin J, Chen G, Xie S, Zeng H, and Chen R. (2011). Colorectal cancer detection by gold nanoparticle based surface-enhanced Raman spectroscopy of blood serum and statistical analysis. *Bio Optics Express, 19(14)*(13565–13577).

Lin D, Pan J, Huang H, Chen G, Qiu S, Shi H, Chen W, Yu Y, Feng S, and Chen R. (2014). Label-free blood plasma test based on surface-enhanced Raman scattering for tumor stages detection in nasopharyngeal cancer. *Sci. Rep, 48*(4751), 324-387.

Lin J, Chen R, Feng S, Pan J, Li B, Chen G, Lin S, Li C, Sun L, Huang Z, and Zeng H. . (2012). Surface enhanced Raman scattering spectroscopy for potential noninvasive nasopharyngeal cancer detection. *Journal of Raman Spectrosc, 43*(4), 497–502.

Macallan C, Noble C, Baldwin C, Foskett M, McManus T, Griffin E. (1993). Prospective analysis of patterns of weight change in stage IV human immunodeficiency virus infection. *Journal of Clinical Nutrition, 58*(3), 24 -417.

Macallan D, Noble C, Baldwin C, Jebb S, Prentice M, Coward W. (1995). Energy expenditure and wasting in Human Immunodeficiency Virus infection. *Journal of Medicine, 333*(2), 8 -83.

Madden, Anne Fischer and Dean. (2011). The Origin and Evolution of HIV. *Student Guide*, 3-7.

Mariani M, Day J, and Deckert V. . (2010). Applications of modern micro-Raman spectroscopy for cell analyses. *Biosens and Bioelectrons, 2*(2-3), 94–101.

Marleen de Veij, Peter Vandenabeele, Krystyn Alter Hall, Facundo M. Fernandez. (2006). Fast detection and identification of counterfeit antimalarial tablets by Raman spectroscopy. *Journal of Raman Spectroscopy, 38*(181-187).

Matthews Q, Jirasek A, Lum J, Duan X, and Brolo A. (2010). Variability in Raman spectra of single human tumor cells cultured in vitro: correlation with cell cycle and culture confluency. *Applied Spectroscopy. 64* (8), 871– 887.

Molly Gregass, Jonathan Scaffid, Benoit Lauly and Tuan. (2010). Surface-Enhanced Raman Scattering Detection and Tracking of Nanoprobes: Enhanced Uptake and Nuclear Targeting in Single Cells. *Applied Spectroscopy, 64(8)*(806), 1254-1269.

Moor, H. Kitamura K. Hashimoto, M. Sawa, B. Andriana, K. Ohtani, T. Yagura, H. Sato. (2013). Study of Virus by Raman Spectroscopy. (D. V. Daniel L. Farkas, Ed.) *Proceedings of SPIE, 8587 85871X-1*, 1-11.

Narayana R. Isola, David L. Stokes, and Tuan Vo-Dinh. (1998). Surface-Enhanced Raman Gene Probe for HIV-1 Detection. *Analitical Chemistry, 70*(7), 1352-1356.

Ning Gan, Xiaowen Du , Yuting Cao, Futao Hu , Tianhua Li and Qianli Jiang. (2013). An Ultrasensitive Electrochemical Immunosensor for HIV p24 Based on Fe3O4@SiO2 Nanomagnetic Probes and Nanogold Colloid-Labeled Enzyme–Antibody Copolymer as Signal Tag. *Materials, 1255-1269*(6), 1-15.

Otero E, Sathaiah S, Silviera L, Pomerantzeff A and Pasqualucci C. (2004). Raman spectroscopy for diagnosis of calcification in human heart valves. *Journal of Applied Spectroscopy, 18*(1), 75–84.

Pascut F, Goh H, George V, Denning C, Notingher I. (2011). Toward label-free Raman-activated cell sorting of cardiomyocytes derived from human embryonic stem cells. *Journal of Biomedical Optics, 16*(045002 – 045002), 1-6.

Pavel Zinin, Ningjie Hu, Lori Kamemoto, Qigui Yu, Anupam Misra and Shiv Sharma. (2011). Raman spectroscopy of HIV-1 antigen and antibody. In E. S. Brian M. Cullum (Ed.),

Proceedings of SPIE. 8025 80250D-1, pp. 1-7. Smart Biomedical and Physiological Sensor Technology.

Pencharz J, Jeejeebhoy K. (1979). Regulation of whole-body protein synthesis. *Lancet, 1(8125):1087.)*, 325-356.

Pencharz P, Parsons H, Motil K, Duffy B. (1981). Total body protein turnover and growth in children: is it a futile cycle? *Medical Hypotheses, 7*(2), 60-155.

Persson. (2003). Smoothing by Savitzky-Golay and Legendre filters. *Math. Systems Theory Biol. Comm. Comp, 134*, 301–316.

Petisco Christina, Balbino Gascia, Criado Inigo and Zabalgogeazoa M. (2011). A spectroscopy approach to the study of virus infection in the endophytic fungus Epichloë festucae. *Virology Journal, 57*(4), 3-21.

Rahim A. Bahmani M, Masoudnejad and Mohammadzade. (2010). Application of near-infrared spectroscopy and support vector machine in detection of HIV-1 infection. *Medical Journal, 25*(3), 1-7.

Raquel Martin, Alison Malkin and Fiona Mary Lyng. (2015). Current Advances in the Application of Raman Spectroscopy fo rMolecular Diagnosis of Cervical Cancer. *BioMed Research International, 2015*, 1-9. doi:doi.org/**10.1155/2015/561242**

Raymond Wai-Yin Sun, Rong Chen, Nancy P. Chung, Chi-Ming Ho, Chen-Lung Steve. (2005). Silver nanoparticles fabricated in Hepes buffer exhibit cytoprotective activities toward HIV-1 infected cells. *Royal Society of Chemistry, 23*(18), 50-56.

RCSB. (2011). *Protein Data Bank*. Retrieved 7 21, 2016, from Structural Biology of HIV-1: http://www.rcsb.org/pdb/home.do1RTD/2XDE/2X7L/2ZD1/3H47/3/JWD/3KLE/3MIA/3 NGB/3OBU/3OS2/3OY9/3PO5

Rosenblatt. (1959). *Principles of neurodynamics* (Vol. **2**). New York: Spartan Press.

Savitzky Golay. (1964). Soothing and differentiation of data by simplified least square. *Analalytical Chemistry, 36* (81), 1627–1639.

Scott and Palmer. (1993). Neural Network Analysis of ventilation perfusion lung scans. *Radiology, 186*, 23-25.

Shafiee H, Asghar W, Inci F, Yuksekkaya M, Jahangir M, Zhang MH, Durmus NG, Gurkan UA, Kuritzkes R, Demirci U. (2015). Paper and flexible substrates as materials for biosensing platforms to detect multiple biotargets. *Scientific reports, 6*(5), 8719.

Shafiee H, Lidstone A, Jahangir M, Inci F, Hanhauser E, Henrich J, Kuritzkes D, Cunningham T, Demirci U. (2014). Nanostructured optical photonic crystal biosensor for HIV viral load measurement. *Scientific reports, 28*(4), 4116.

Shafiee H, Wang S, Inci F, Toy M, Henrich J, Kuritzkes R, and Demirci U . (2015). Emerging technologies for point-of-care management of HIV infection. *Annual review of medicine, 66*(7), 387–405 .

Smith J, Dent Geoffrey and Edwin B. (2005). *Modern Raman Spectroscopy -A Practical Approach.* Chichester: John Wiley & Sons Ltd.

Smith Lindsay. (2002). A tutorial on Principal Components. *Statistical Analysis, 45*(4), 586-601.

Stekler J, Maenza J, Stevens CE, Swenson PD, Coombs RW, Wood RW, Campbell MS, Nickle DC, Collier AC and Golden MR. (2007). Screening for acute HIV infection: lessons learned. *Clinical Infectious Diseases, 44*(7), 459–461.

Thermo Electron. (2003). *Introduction to Raman Spectroscopy.* New York Press.

Timothy. (2010). Practical Applications of Surface-Enhanced Raman Scattering. *Thermo Fisher Scientific, 51874*.

Urmos J, Sharma K, and Mackenzie T. (1991). Characterization of some biogenic carbonates with Raman spectroscopy. *Am. Mineral, 76*(3), 641–646.

Virkler K, Lednev K. (2009). Analysis of body fluids for forensic purposes: from laboratory testing to non-destructive rapid confirmatory identification at a crime scene. . *Forensic Science International, 188*(3), 1-17.

Wold Svante. (1995). Chemometrics; What do we mean with it, and waht do want from it? *Chemometrics and Intelligent Laboratory Systems, 30* (1995), 109-115.

World Health Organization (2017).*World Health Organization*. [online] Available at: http://www.who.int/hiv/en/ [Accessed 23 Aug. 2017].

Wu Y, Doi K, Giger L, Vyborny J, Schmidt R, Nishikawa R. (1991). Application of artificial neural networks in mammography for the detection of breast cancer. *Radiology, 181*, 143-154.

Wu Y, Giger ML, Doi K, Vyborny CJ, Schmidt RA, Metz CE. (1993). Artificial neural networks in mammography: application to decision making in the diagnosis of breast cancer. *Radiology, 187*(7), 187.

Ying-Sing, Jeffrey and S. Church. (2014). Raman spectroscopy in the analysis of food and pharmaceutical nanomaterials. *journal of food and drug analysis, 22*(29-48).

Zaman Grace Wu and Muhammad H. (2012). Low-cost tools for diagnosing and monitoring HIV infection in low-resource settings. *Bulletin of the World Health Organization, 45*(3-5), 3-9.

Zhou L, Huang, J Yu, Liu Y and You T. (2015). A Novel Electrochemiluminescence Immunosensor for the Analysis of HIV-1 p24 Antigen Based on P-RGO@ Au@ Ru-SiO2 Composite. *ACS applied materials & interfaces, 7*(44), 24438-24445.

Appendix A: Informed Consent Form

Introduction: There are two types of HIV which can infect humans (HIV-1 and HIV-2). Both types are transmitted in the same way and they appear to cause clinically indistinguishable AIDS. However, HIV-2 is less virulent and not easily transmitted. It is mainly found in western parts of Africa, and the period between initial infection and illness is longer than in the case of HIV-1. HIV-1 is the main cause of majority of HIV infections globally. There are three subgroups of HIV-1: HIV-1-M, HIV-1-N and HIV-1-O. HIV-1-M is the most common and has spread world wide. At the beginning of HIV infection, the virus antigen called p24 dominates and can be detected within few days after infection. As the body develops a counteractive mechanism to fight the virus, HIV antibodies are produced. The time taken to detect these antibodies may vary due to a number of factors but not earlier than three months. HIV infection gradually destroys the body's defense cells used in fighting infections, leaving the body susceptible to diseases it would normally be able to fight. When the virus is not diagnosed in time and proper monitoring mechanism is not put in place, the immune system becomes weaker, and the person with HIV will begin to develop opportunistic infections. At this point the person with HIV will begin to develop AIDS. This study will be investigating how the sensitivity of Confocal Raman Microscope can be applied in the rapid and early detection of HIV-1 and also its quantification based on the corresponding changes of Raman intensities from different viral load.

The Purpose of the Study: To develop a rapid and sensitive real time technique that should simultaneously elucidate traces and quantification of HIV-1 in blood/serum through chemometric Raman spectral profile. The unique spectral profile of p24 antigen will be used as reference spectra in classifying blood samples and quantifying the infected blood samples.

Procedure: The study will require you to donate blood once. During this donation, you may be required to undergo a voluntary HIV test. If you agree to donate, you will be required to sign two copies of Informed Consent Form confirming that you have been informed about the study and voluntarily agree to take part. One copy is yours to keep and the other copy will be kept in our confidential study file. If you do not wish to keep your copy, you will sign a form that states that you do not want to take it, and we will keep it for you. You will be asked questions about your general health and a medical examination will be performed. A 10 ml (about 1 tablespoon) of

your blood will be drawn for laboratory HIV tests (ELISA test) as well as sample for Raman fingerprinting.

What are the Risks and /or Discomforts: There may be some risk involved in drawing blood for the laboratory test such as: You may have pain and bruising where the needle goes into your arm. You may feel dizzy or faint. There are no anticipated physical risks in participating in this study. However, if there are any injuries that may arise due to your participation, you will be offered free treatment.

What are the benefits of study participation: There are no direct benefits to you. However, the outcome of this study may help in molecular virologist to speed up identification of HIV-1. It will also reduce the time taken to detect the presence of the virus after a recent HIV-1 exposure.

Injuries: We do not expect you to be injured as a result of being in this study. If you are hurt as a result of being in this study, the clinic stuff will give you the necessary treatments for the injuries including emergency treatment for free. You will not receive any money or other forms of payment for such injuries. You will not give up any legal rights by signing this Consent Form.

When can you leave the study: Your participation in this study is completely voluntary, you can leave this study at any time without giving a reason. Withdrawal will NOT compromise any rights you had before entry into the study or influence any current or future medical care you may need. If you leave or are asked to leave the study after lab tests have been done, you may still get your test result from the donation center.

Confidentiality: Your participation in the study, all information collected about you, and all laboratory test results will be available to no one except the study team. You will only be identified only by your unique identity number, which is known only by you and the clinic staff. Your identity will not be disclosed in any publication or presentation of this study.

Storage of samples: The samples will be stored for four months there after they will be destroyed and disposed off. If you don't want your samples to be stored we give you a period of one month to contact us. If you will not come, we will assume that you have agreed that your samples be stored for further analysis.

Please tick one of the following.

☐ I don't want my samples to be stored

☐ I agree that my samples be stored for further analysis.

Contact Number: If you have any question regarding the study or your participation in the study, you can call **Ben Otange,** the principle investigator, **mobile no: 0729901867.** If you have a question about your rights as a volunteer you should contact **Prof. Chindia**, the chairman of Ethics Committee at Kenyatta National Hospital, **Tell: 726300-9, Fax: 725272**. You can also contact **Dr. J Oyugi** (medical doctor and a supervisor) on **0713898564.**

Appendix B: Informed Consent Document

I, (name of the volunteer)………………………………………………………………………..of (address) …………………………………………………………………………agree to take part in this research project entitled: **"Early and Rapid Detection of Human Immunodeficiency Virus Using Vibrational Raman Spectroscopy"** I have been told in details about the study and know what is required of me. I understand and accept the requirements. I understand that my consent is entirely voluntary and that I may withdraw from the research study for any reason, and this will not affect the legal rights I may otherwise have. My questions have been answered to my satisfaction.

Participant: Print name……………………………………………………………………

Signature/ Mark or Thumbprint: …………………………………………………………

Date: ………………………………………………………………………………..

Person Obtaining Consent

I have explained the nature, demands and foreseeable risks of the above study to the volunteer and answered his/her questions:

Print Name: …………………………………………………………………………

Signature/ Mark or Thumbprint: …………………………………………………………

Date: ………………………………………………………………………………..

Impartial Witness: (Only necessary if volunteer was not able to read and understand the Consent Information Form and Informed Consent Document):

I affirm that the informed consent Document has been read to the volunteer and he/she understand the study, had his/her questions answered and I have witnessed the volunteer's consent to study participation.

Print Name: …………………………………………………………………………

Signature/ Mark or Thumbprint: …………………………………………………………

Date: ………………………………………………………………………………..

Appendix C: Consent form and Document Translated in Kiswahili

Fomu ya Ruhusa Halali ya Mshiriki

Utangulizi: Kuna aina mbili za virusi vya ukimwi ambazo zinawambukiza binadamu nazo ni; Virusi vya ukimwi-1 na virusi vya ukimwi-2 ukipenda (HIV-1 na HIV-2). Aina hizi mbili hupitishwa kwa njia sawa na huonekana kusababisha UKIMWI usiokuwa na tofauti. Hata hivyo, virusi vya ukimwi-2 huwa si ya kivirusi sana na huwa haipitishwi na kuambukizwa kiurahisi. Aina hii hupatikana sana katika maeneo ya Magharibi ya Afrika na kipindi chake cha uambukizo huwa mrefu ukilinganisha na ile aina ya Virusi vya ukimwi-1. Virusi vya ukimwi-1 ndiyo huwa kisababishi kikuu cha mengi ya maradhi ya kivirusi ulimwenguni.

Kunazo vikundi vitatu vya Virusi vya ukimwi-1ambavyo ni kama ifuatavyo: HIV-1-M, HIV-1-N na HIV-1-O. Aina hii ya HIV-1-M ndiyo iliyo ya kawaida sana na imeenea ulimwengu mzima. Katika kuanza kwake kwa uambukizi, Virusi antigeni iitwayo p24 huchangia na yaweza kubainika au kutambulika siku chache baada tu ya kuambukizwa.

Namna mwili unavyojenga njia murwa nay a haraka kupigana ili kushinda virusi, wakati huo huo virusi antibody huzalishwa. Muda unauchukuliwa ili kutambua antibody hizi, huenda ukatofautiana kutokana na ambo kadha wa kadhahata hivyo siyo chini ya miezi mitatu.

Maambukizi vya ukimwi kwa muda huharibu kinga ya mwili inayotumika katika kupigana na maambukizi mwilini. Jambo linaloacha mwili na asilimia ndogo ya kuweza kupigana na magonjwa kadha wa kadha amabyo kwa kawaidamwili hupigana nao.

Iwapo virusi havijabainika au kutambulika kwa wakati ufaao na utunzi unaofaa usipozingatiwa, kinga ya mwili huwa dhaifu na mtu aliye na virusi hivi huanza kujenga nafasi nyingi za kuambukizwa maradhi. Katika kiwango hiki, mtu aliye na virusi huanza kupata UKIMWI. Kwa hivyo, utafiti huu utakua ukitaii na kubaini ni vipi 'Confocal Raman Microscope' yawezatumika kutambua na kubaini kwa mapema kuwepo kwa virusi na pia kiwango chake kwa kuzingatia tofauti zinazolingana na pateni za Raman zinazotokana na uzito wa virusi

Sababu Kuu ya Utafiti: Ni kujenga na kubuni njia ya harakana ya kiwakati tena ya kipekeeambayo yafaa kutumika kimfululizo ili kusababisha na kutambua viwango vya virusi-1 katika damu au mbegu za uzazi kwa kupitia chemometric pateni za Raman ya p24 ambayo ni

antigeni ya kipekee itakayo tumika kama rejeleo la kuweka sampuli za damu kwa makundi na pia kutambua na kubaini viwango vya maambukizi katika sampuli za damu.

Utaratibu: Utafiti huu utakuitaji kutoa damu mara mojakatika kipindi hiki cha utoaji wa damu waweza hitajika kupitia wasia wa kujitolea na pia kumimwa virusi vya UKIMWI. Kama utakubali kutoa damu, utaitajika kutia sahihi nakala mbili za ruhusa halali ya mshirika ili kuthibitisha kuwa umeambiwa kuhusiana na utafiti huu na kujikubali mwenyewe kuhusika katika utafiti huu. Nakala moja ni yako uiweke na nyingine itawekwa kwa faili yetu ya usiri. Kama hutaki kuiweka nakala yako ambayo umetia sahihi utaitajika kutia sahihi fomu ya kuonyesha kuwa hutaki kuichukua fomu hiyo na tutakuwekea fomu hiyo. Pia utaulizwa maswali kadha kuhusiana na afya yako ya kijumla na pia utafiti wa kimatibabu utafanyiwa. Mililita kumi (10 ml) ya damu ambayo ni kama kijiko cha chai utachukuliwa ili kujaribiwa katika mahabara pia ELISA test au vipimo vya ELISA vitafanyika, vilevile sampuli za patani za Raman.

Tahadhari na / ama ubaya wake: Hizi huenda zikawa baadhi ya athari ambazo huenda zikapatikana katika kuchukulia kwa damu ili ijaribiwe katika mahabara. Baadhi huenda zikawa kuwepo kwa uchungu kutokana na kudungwa shindano. Hii huenda likasababisha uchungu au kuwepo kwa kuzirai. Hata hivyo hamna majerui yanayo taarajiwa kutokana na kushiriki katika utafiti huu na hivyo iwapo patatokea majiraha yeyote kutokana na kushiriki kwako katika utafiti hii basi utapea matibabu ya bure.

Manufaa ya Kushiriki katika Utafiti huu: Huenda hamna faida za moja kwa moja utakazozipata baada ya kushiriki hata hivyo matokeo ya utafiti huu huenda yakasaidia utafiti wa virusi ambavyo vitasaidia kuongeza kasi katika kutambua kwa virusi-1au HIV-1 pia hupunguza muda unaochukuliwa katika kutambua na kubaini kuwepo kwa virusi baada ya kuwa na virusi-1 ama HIV-1 katika siku au muda uliopita

Majeraha: Kwanza hatutarajii wewe kuumia kutokana na kushiriki katika utafiti huu. Iwapo utaumia kutokana na kushiriki utafiti huu, maafisa wa zahanati watakupa matibabu yanayohitajika yakiwemo pia ya dharura na ya bure kulingana na majeraha hayo. Hata hivyo, hutapokea pesa au hata malipo ya aina nyingine ili kufidia majeraha. Hutatupilia haki zako za kimsingi kutokana na kutia sahihi fomu hii ya ruhusa halali kutoka kwako.

Usiri / Ufaragha: Kushiriki kwako katika utafiti huu, habari yote itakayokusanywa kutoka kwako na majaribio na matokeo ya mahabara haitafunguliwa wala kuwekwa wazi kwa yeyote isipokuwa kwa kikosi cha utafiti. Utatambulishwa tu kwa nambari yako ya kipekee ambayo inajulikana nawe tu pamoja na wataalam wa zahanati. Jina lako halitatajwa wala kuwekwa wazi au hata kuchapishwa wala kuwasilishwa katika utafiti huu.

Kuwekwa kwa Sampuli: Sampuli hii itawekwa kwa miezi mine hatimaye itaharibiwa na kuupwa. Na iwapo hutaki sampuli yako iwekwe nasi, tutakupa muda wa mwezi moja uwasiliane nasi. Iwapo hutakujia sampuli yako baada ya mawasiliano tutakisia kuwa umekubali kuwa sampuli yako iwekwe na itumike katika utafiti Zaidi. Tafadhali weka alama katika moja wapo ya haya

- ☐ Sitaki sampuli yangu ihifadhiwe au kuwekwa
- ☐ Nakubali kuwa sampuli yangu iwekwe kwa utafiti Zaidi

Nambari ya Mawasiliano: Iwapo utakuwa na swali lolote kuhusiana na utafiti huu, waweza kuwasiliana na **Ben Otange** mtafiti mkuu, Nambari ya Rununu ni **+254729901867**. Na kama unalo swali lolote kuhusiana na haki zako kama mtu wa kujitolea wafaa uwasiliane na **Prof. Chindia** ambaye ni mwenyekiti wa kamati ya maadili katika Hospitali ya Kitaifa ya Kenyatta kwenye nambari **726300-9** au faksi **725272**. Pia waweza wasiliana naye **Dr. J Oyugi** (Daktari wa dawa na mwelekezi) kwa nambari hii **+254713898564**

Dokumenti ya Ruhusa Halali ya Mshirika

Mimi, (Jina lako) …………………………………………..……………………………wa (anwani yako) …………………………………………………… nakubali kushirika katika utafiti huu wenye mada **"Njia ya haraka na rahisi ya kubaini na kutambua virusi kwa kutumia Raman"** au katika kimombo **"Early and Rapid Detection of Human Immunodeficiency Virus Using Vibrational Raman Spectroscopy"**

Nimeambiwa kwa mapema kuhusiana na utafiti huu na sasa najua ninalotarajiwa kulifanya. Naelewa na nakubali yanayohitajika. Naelewa kwamba kukubali kwangu ni kwa kujitolea tu na kuwa naweza kujitoa kwa utafiti huu kutokana na sababu yeyote na hili halitadhuru haki zangu za kimsingi ambazo naweza kuwa nazo. Naamini maswali yangu yamejibiwa vilivyo hadi kwa kiwango cha utoshelezi wangu.

Jina la Mshiriki …………………………………………………………………………

Sahihi/Alama/Alama ya kidole cha gumba ………………………………………….

Terehe …………………………………………………………………………………..

Kwa Mtu Apataye Ruhusa Halali

Nimeeleza na kufafanua mahitaji na athari ambazo huenda zikatokea kwa mshiriki wa kujitolea kutokana na utafati na hata kuyajibu maswali yake.

Jina …………………………………………………………………………………………

Sahihi/Alama/Alama ya kidole cha gumba ………………………………………….

Terehe …………………………………………………………………………………..

Shahidi asiye kamilifu: (ni muhimu tu iwapo aliyejitolea hakuwa na uwezo wa kusoma na kuelewa linalopatikana katika fomu na pia katika nakala ya mshiriki wa ruhusa yake halali)

Nathibitisha kuwa fomu ya kutaka na kukubali imesomwa mbele ya mshiriki wa kujitolea na yeye ameelewa utafiti na maswali yake yamejibiwa vilivyo na hivyo mimi nimeshuhudia ruhusa ya kujitolea kushiriki utafiti.

Jina ………………………………………………………………………………….

Sahihi/Alama/Alama ya kidole cha gumba ………………………………………….

Terehe …………………………………………………………………………………..

Appendix D: Bioethical Approval by KNH-UoN Ethics and Research Committee

UNIVERSITY OF NAIROBI
COLLEGE OF HEALTH SCIENCES
P O BOX 19676 Code 00202
Telegrams: varsity
(254-020) 2726300 Ext 44355

KNH-UON ERC
Email: uonknh_erc@uonbi.ac.ke
Website: http://www.erc.uonbi.ac.ke
Facebook: https://www.facebook.com/uonknh.erc
Twitter: @UONKNH_ERC https://twitter.com/UONKNH_ERC

KENYATTA NATIONAL HOSPITAL
P O BOX 20723 Code 00202
Tel: 726300-9
Fax: 725272
Telegrams: MEDSUP, Nairobi

Ref: KNH-ERC/A/16

13th January 2016

Ben Otieno Otange
(SM13/23519/14)
Egerton University
P O Box 536-20115
EGERTON

Dear Mr. Otange

Revised research proposal: Investigation of Early Detection of Human Immunodeficiency Virus using Raman Spectroscopy (P637/10/2015)

This is to inform you that the KNH- UoN Ethics & Research Committee (KNH-UoN ERC) has reviewed and **approved** your above proposal. The approval periods are 13th January 2016 –12th January 2017.

This approval is subject to compliance with the following requirements:

a) Only approved documents (informed consents, study instruments, advertising materials etc) will be used.
b) All changes (amendments, deviations, violations etc) are submitted for review and approval by KNH-UoN ERC before implementation.
c) Death and life threatening problems and serious adverse events (SAEs) or unexpected adverse events whether related or unrelated to the study must be reported to the KNH-UoN ERC within 72 hours of notification.
d) Any changes, anticipated or otherwise that may increase the risks or affect safety or welfare of study participants and others or affect the integrity of the research must be reported to KNH- UoN ERC within 72 hours.
e) Submission of a request for renewal of approval at least 60 days prior to expiry of the approval period. (*Attach a comprehensive progress report to support the renewal*).
f) Clearance for export of biological specimens must be obtained from KNH- UoN ERC for each batch of shipment.
g) Submission of an *executive summary* report within 90 days upon completion of the study.
   This information will form part of the data base that will be consulted in future when processing related research studies so as to minimize chances of study duplication and/ or plagiarism.

For more details consult the KNH- UoN ERC website **http://www.erc.uonbi.ac.ke**

Protect to discover

Appendix E: Safety Precautions Taken When Handling p24 antigen

This preparation should be regarded as potentially hazardous to health. It should be used and discarded according to laboratory's safety procedures. Such safety procedures include;

i)   Wearing of protective gloves and avoiding the generation of aerosols.

ii)   Take care when opening p24 antigen ampoule to avoid cuts.

iii)   Take care that no material is lost from the ampoule and that no glass falls into the ampoule.

iv)   Unopened ampoules should be stored at -20^0C

v)   Incase the content comes in contact with the eye or skin wash thoroughly with too much water.

vi)   Incase of either inhalation or ingestion seek quick medical advice.

vii)   Spillage of ampoule contents should be taken up with absorbent material wetted with an appropriate disinfectant. Rinse area with an appropriate disinfectant followed by water. Absorbent materials used to treat spillage should be treated as biological waste. Alternatively, the contents may be laced in a hot water bath with temperature 56^0C or more for a period one hour before disposal.

viii)   Handle all reagents and samples as if capable of transmitting disease. The Conjugate Solution contains human derived material, and the HIV-1 p24 Standard contains human and virus derived materials. Although these reagents have been inactivated, there is no absolute assurance that such products cannot transmit infection.

Appendix F: Plates

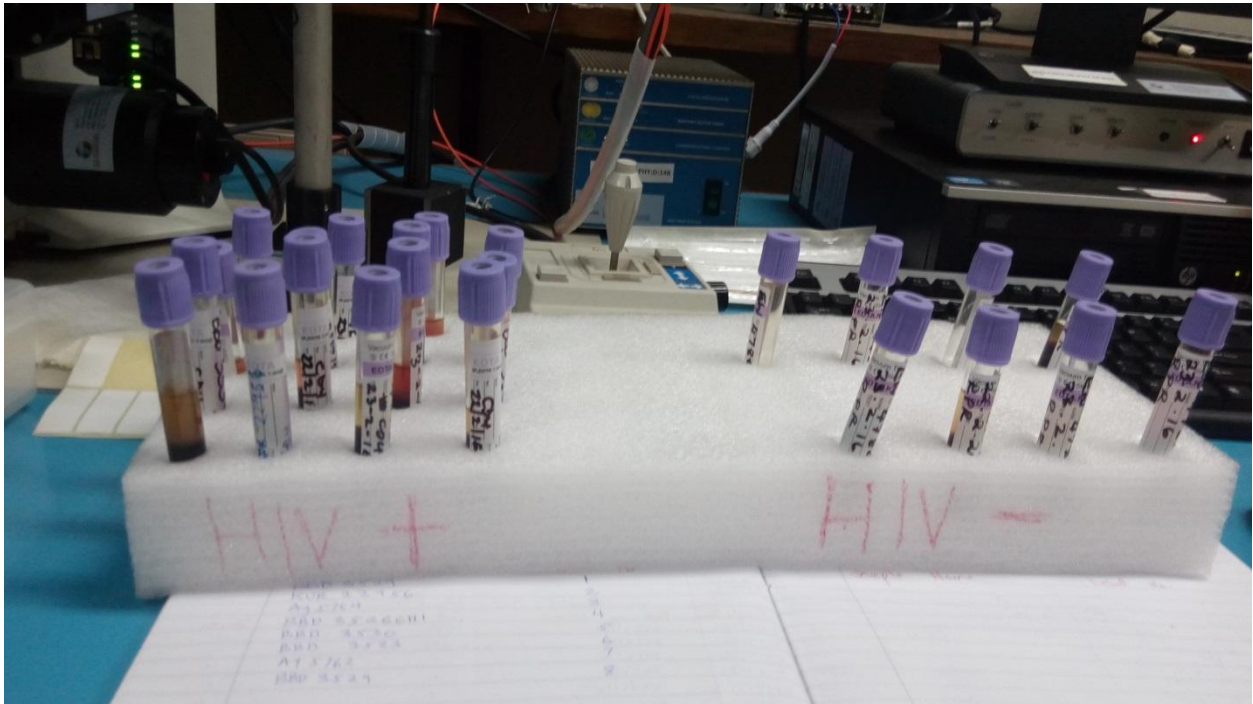

Plate 1: Samples placed in a vertical position to aid plasma and blood separation.

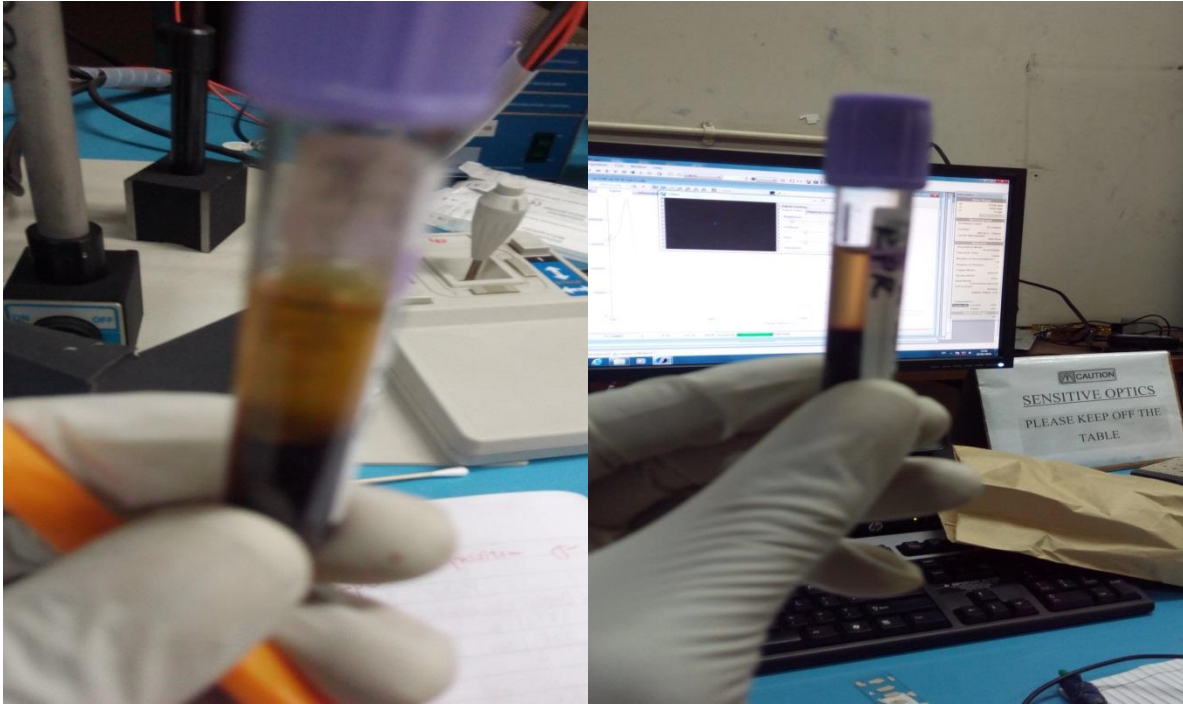Plate 2: p24 antigen within unopened ampoule

Plate 3: EDTA collection tube with blood and plasma separated

Plate 4: Blood samples adsorbed on Ag substrate on a microscope glass slide

Plate 5: Confocal Raman Microscope

Plate 6: Raman Optical Filters

Appendix G: List of Publications

i) **Otange, B**, Rop, R., Oyugi, Julius. O, and Birech, Z. (2016, October). Rapid detection of HIV1-p24 antigen in human blood plasma using Raman spectroscopy. In *Frontiers in Optics* (pp. FF5A-5). Optical Society of America. https://www.osapublishing.org/abstract.cfm?URI=FiO-2016-FF5A.5

ii) **Ben O. Otange**, Zephania Birech, Justus Okonda and Ronald Rop (2017). Conductive silver paste smeared glass as Raman sample substrates for label-free detection of HIV-1 and HIV-1 p24 antigen in blood plasma. Journal of analytical and bioanalytical chemistry *DOI 10.1007/s00216-017-0267-0.* Springer

Appendix H: List of Conference/Seminar Presentations

i) *"Multivariate analysis as a way of discriminating blood and plasma Raman dataset for HIV-1 infection"* in the **First Egerton University Physics Department Seminar.** Physics Department Conference room on 29th June 2016

ii) *"Rapid detection of HIV-1 p24-antigen in human blood plasma"* **Frontiers in Optics/Laser Science Conference (FiO/LS)** held on 17-21 October 2016 at Rochester Riverside Convention Center, New York

iii) *"Surface Enhanced Raman Spectroscopic Characterization of blood and plasma samples for HIV-1 infection"* in the **11th JKUAT Scientific, Technological and Conference and Exhibitions.** JKUAT University from 10-11th November 2016

iv) *"Label-free Surface Enhanced Raman Spectroscopic Detection HIV-1 Infection in blood and plasma adsorbed on conductive silver pasted glass substrate."* in the **11th Egerton University International Conference.** Egerton University from 30th - 1th March 2017.

v) *"Early Detection And Concentration Determination Of HIV-1 Based on The Associated Surface Enhanced Raman Spectroscopic Peaks of Infected Human Plasma"* **In The 3rd Africa International Biotechnology and Biomedical Conference (AIBBC 2017).** SAJOREC JKUAT University from 12th -16th Sep 2017