Development of Raman Spectroscopy as a Clinical Diagnostic Tool

by

Santa Borel

A thesis submitted in conformity with the requirements
for the degree of Master of Science

Department of Medical Biophysics
University of Toronto

Development of Raman Spectroscopy as a Clinical Diagnostic Tool

Santa Borel

Master of Science

Department of Medical Biophysics
University of Toronto

2016

Abstract

Raman spectroscopy is the collection of inelastically scattered light in which the spectra contain biochemical information of the probed cells or tissue. This work presents both targeted and untargeted ways that the technique can be exploited in biological samples. First, surface enhanced Raman scattering (SERS) gold nanoparticles conjugated to targeting antibodies were shown to be successful for multiplexed detection of overexpressed surface antigens in lung cancer cell lines. Further work will need to optimize the conjugation technique to preserve the strong binding affinity of the antibodies. Second, untargeted Raman microspectroscopy combined with multivariate statistical analysis was able to successfully differentiate mouse ovarian surface epithelial (MOSE) cells and spontaneously transformed ovarian surface epithelial (STOSE) cells with high accuracy. The differences between the two groups were associated with increased nucleic acid content in the STOSE cells. This shows potential for single cell detection of ovarian cancer.

Acknowledgments

I would like to thank my supervisor Dr. Brian Wilson for his support in my project and for providing me with so many different opportunities for growth and learning. He has taught me that every problem has a solution and to always continue searching. I would also like to thank my committee members Drs. Gang Zheng, Richard Hill and Thomas Kislinger for their interest in my work and constructive feedback and discussions during our meetings.

I would like to thank Dr. Patrick McVeigh for beginning and establishing the SERS project. His endless support and encouragement helped me develop my problem solving skills and stay motivated.

I would also like to thank Dr. Sangeeta Murugkar for presenting me with such a stimulating project working with intrinsic Raman spectra of cells. I was able to develop my ideas further and have a better understanding of the field. Her optimism and excitement for the project made it a pleasure to work with her. I also acknowledge the contributions of Dr. Barbara Vanderhyden and Emil Prikryl to the success of the work.

James Jonkman and the AOMF staff were a great help in supporting and training me on the Raman microscope and answering any of my microscopy questions. I appreciate the help from my fellow biophotonics lab members, specifically Carl Fisher, Carolyn Niu and Sam Lim. I enjoyed all the time I spent with my fellow MBP students and I value the close friendship that I was able to make.

I appreciate all the love and support I was given by my family and friends that has helped guide me throughout this whole research project and Master's program.

Table of Contents

List of Tables

List of Symbols and Abbreviations

Au	Gold
BSA	Bovine serum albumin
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
EGFR	Epidermal Growth Factor Receptor
EpCAM	Epithelial cell adhesion molecule
HER2	Human epidermal growth factor receptor 2
HGSC	high-grade serous ovarian cancer
IgG	Immunoglobulin G
LDA	Linear discriminant analysis
LOOCV	Leave-one-out-classification validation
MAL	maleimide
MESA	2-mercaptoethanesulfonic acid
MOPS	(N-morpholino) propanesulfonic acid
MOSE	Mouse ovarian surface epithelium
NHS	N-hydroxysuccinimide
NPs	Nanoparticles
PBS	Phosphate-buffered saline

PC	Principal component
PCA	Principal component analysis
PEG	Polyethylene Glycol
RNA	Ribonucleic acid
SERS	Surface enhanced Raman scattering
STOSE	Spontaneously transformed surface epithelium

List of Figures

Chapter 1
Introduction

1.1. Motivation

Cancer is the leading cause of death in Canada, accounting for 30% of all deaths [1]. Important research has gone into understanding the molecular origins of the disease and improving the ways that this can be used to detect it earlier, delineate tumor margins and determine the optimal treatment plan.

Cancer is usually diagnosed using a macroscale imaging modality and confirmed through histopathology. Diagnosis imaging modalities such as computed tomography (CT), magnetic resonance imaging (MRI), positron emission tomography (PET) and ultrasound are often the first methods. These techniques can provide information about the tumor such as is 3D volume, structure and composition. A suspicious lesion is then confirmed by biopsy, followed by histological analysis. Tissue is extracted from the patient from a suspicious region, then sliced, mounted on a slide and stained. A trained histopathologist can determine whether the cells seen in this small piece of tissue are normal or malignant. This process is very time consuming and can delay a diagnosis or increase surgery time when it is used to determine tumor margins such as in Mohs micrographic surgery. While these methods are invaluable for cancer diagnosis, there are challenges of incorporating these techniques intra-operatively, obtaining high resolution images, obtaining biochemical information, and using multiple probes simultaneously for targeting. This has led to the development of Raman techniques to improve tissue diagnosis [2].

Raman techniques can allow for close to real time biochemical analysis of the tissue saving time and allowing for information to be extracted *in situ* without tissue extraction. The introduction will cover the fundamentals of Raman spectroscopy, followed by a chapter describing surface enhanced Raman scattering (SERS), with the application of optical probes for molecular diagnosis of cancer. Then the third chapter will describe the intrinsic Raman spectra of biological cells and tissue and how malignancy-associated changes can be detected. The final chapter will discuss future work and translating Raman techniques into clinical practice.

1.2. Raman Spectroscopy

1.2.1. Light Interactions with matter

Photons can be scattered when interacting with matter in two ways. Most commonly, this happens in an elastic (Rayleigh) manner, where the scattered photon has the same energy, $h\upsilon$, as the incident wavelength, $h\upsilon$. In a rare scattering event, the scattered photon will have a different energy, $h\upsilon \pm \Delta E$, from the incident photon. This can be visualized using a Jablonski diagram, **Figure 1.1**. From the quantum mechanical interpretation of these scattering events, when an incident photon interacts with a molecule, the molecule is excited to a virtual energy state. From here there are three different potential outcomes. The molecule can relax elastically back to the ground state and emit a photon of equal energy (Rayleigh scattering), $h\upsilon$. In the second type of interaction the molecule can relax from the virtual state to a real phonon state (vibrational mode) emitting a photon with less energy than the incident photon, $h\upsilon - \Delta E$; this is called Stokes shifted Raman scattering. The third possibility is that the photon interacts with a molecule that is already in a higher phonon state and is excited to a higher virtual state and relaxes back down to the ground state with the scattered photon having more energy, $h\upsilon + \Delta E$. This is called anti-Stokes Raman scattering. This type of Raman scattering is less likely to occur, since most molecules are found in a ground state at room temperature and, therefore, undergo Stokes Raman scattering. Due to this, the Stokes Raman spectra are more commonly used in biological specimen analysis.

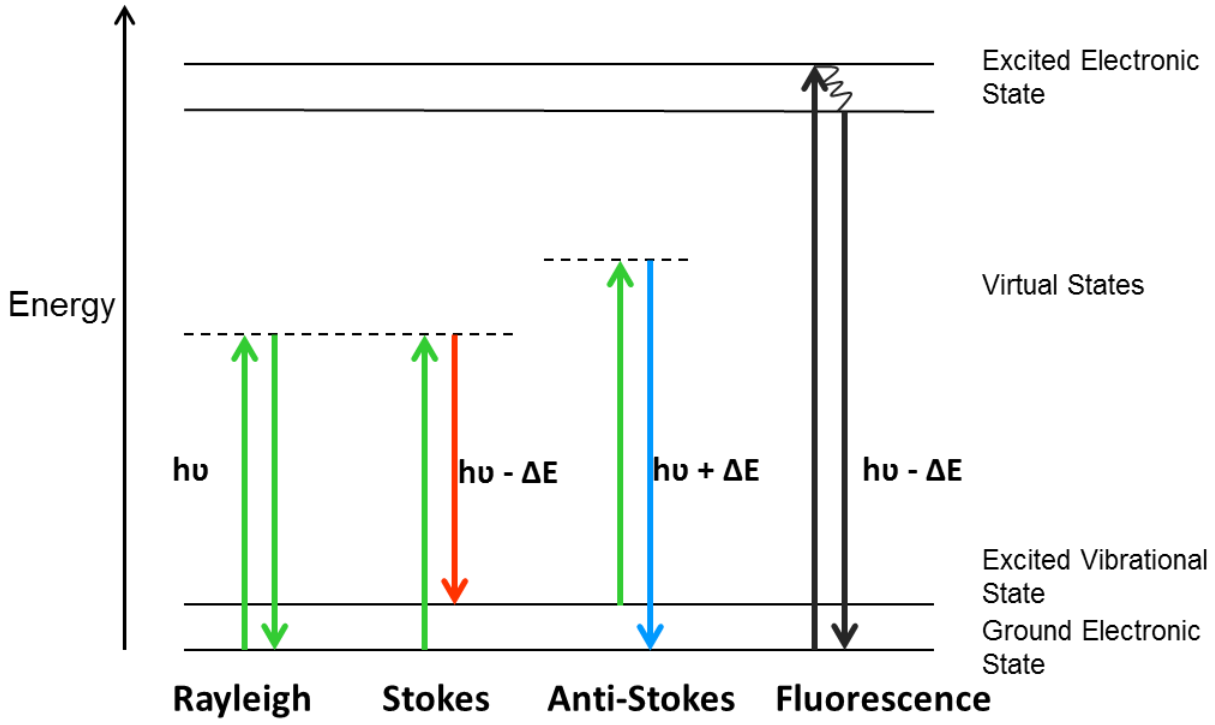

Figure 1.1 Jablonski diagram detailing Rayleigh and Raman (Stokes and anti-Stokes) scattering events and a fluorescence event

Comparing Raman scattering to fluorescence, fluorescence is a very different optical process. In a fluorescence process, the photon is absorbed from the ground singlet state, S_0, to the first singlet state, S_1. At this point the molecule will undergo a vibrational relaxation process until it reaches the vibrational ground state of S_1. It will remain at this energy level for a few nanoseconds, known as the molecule's fluorescence lifetime, and relax back to the ground singlet state, S_0, releasing a photon. Fluorescence differs from the Raman scattering process in that the emission process in fluorescence is independent of the initial photon energy as long as this is sufficient to excite the higher electronic state [3].

All molecules are composed of atoms that are connected to each other by bonds. At above 0 K, molecular bonds vibrate with a certain energy that is dependent on the type of atoms involved, the strength of the bond, and the arrangement of the bond in 3D space, which is partly dependent on the environment. Thus, molecules have a vibrational 'fingerprint' that is characteristic of the bonds that it contains. Using monochromatic light, this 'fingerprint' can be probed by observing

the interaction of the bonds with the oscillating electromagnetic field with energy, E. This induces a polarization of the molecule, given a molecule's polarizability, α,

$$\mu_{ind} = \alpha \tag{1.1}$$

Not all bonds are 'Raman active' as there are selection rules. The vibration of the bond must cause a change in the polarizability of the molecule,

$$\left(\frac{d\alpha}{dq}\right) \neq 0 \tag{1.2}$$

, where q is the simplified coordinate describing the vibrational motion around the equilibrium. Thus, an atom which has an isotropic polarizability, where the polarizability does not change with vibration, will just re-radiate the photon at the incident energy (Rayleigh scattering). A molecule with anisotropic polarizability of the bond will give rise to Raman scattering. [4]

Molecules probed with monochromatic laser light will scatter the incident light revealing information about the bonds contained in them if the bonds are 'Raman active'. The collected Raman spectra record the amplitude of the frequency-shifted light. The spectra are usually represented by the intensity *vs.* the Raman shift, which is defined by the wavenumber, ω,

$$\omega = \left(\frac{1}{\lambda_0} - \frac{1}{\lambda_1}\right) \tag{1.3}$$

, where λ_0 is the incident wavelength and λ_1 is the scattered wavelength. Wavenumber is independent of the incident wavelength and represents the energy difference between the excited and the ground vibrational energy levels of the molecule.

1.2.2. Applications

Spectroscopy is the study of the interaction of light with matter evaluated as a function of wavelength of the emitted or scattered light. Examining the Raman scattering spectrum of a

substance is a rigorous, established spectroscopic technique that has been used for quantitative analysis of molecular materials for many years. Every compound in a sample gives a unique Raman spectrum that is related to the vibrational modes of the molecules. Since Raman scattering is a linear process, a mixture of different substances gives a superposition of each of its constituent Raman spectra allowing for a quantitative analysis of the components.

Raman spectroscopy has grown to be useful in many fields, including pharmaceuticals, environmental sciences, semiconductors, archeology and forensic sciences. For example, in pharmaceuticals, it is used in almost every stage of drug discovery and development from pre-formulation to formulation analytics to quality control [5]. It is a versatile tool that can be used on substances in many forms such as transparent, opaque, colored and also solids, suspensions, and solutions [6]. It is used in DNA analysis to observe oligonucleotides and mononucleotides [7,8].

1.3. Surface Enhanced Raman Scattering

Surface enhanced Raman scattering was first reported in 1974 by McQuillan, who observed that pyridine adsorbed to a roughened silver electrode caused an enhanced Raman signal. [9] This initial article has been cited over 3000 times and this effect has been exploited in a variety of settings.

1.3.1. Physics

Generally, the intrinsic Raman signal of a substance is very weak. The Raman signal of a scatterer can be enhanced through the surface enhanced Raman scattering (SERS) effect. Oscillating free electrons, or surface plasmon, in a metallic material interact with an electromagnetic field which causes an enhancement of the electric field immediately surrounding the surface due to surface plasmon resonance. This is called the Local Field Intensity Enhancement Factor (LFIEF) which is the ratio between the intensity of the electric field in the absence of the metal compared to the intensity of the field in the presence of the metal. The LFIEF is dependent on the geometry and size of the metal structure indicated by the distance from the surface term, r, and the frequency of the electromagnetic field, ω.

$$LFIEF(r,\omega) = \frac{|E(r,\omega)|^2}{|E_0(r,\omega)|^2} \quad (1.4)$$

The Raman process will experience an enhancement factor (EF), from both an incoming photon from the laser, ω_L, and the scattered photon, ω_S. It can be approximated, ignoring any polarization effects, by [10]

$$EF = LFIEF(\omega_L) \times LFIEF(\omega_S) \quad (1.5)$$

The metal substrate for SERS needs to have nano-sized features for the surface plasmon interaction with the field to have a significant enhancement effect, **Figure 1.2**. The most optimal enhancement of the Raman signal occurs when the molecule is adsorbed to the surface (minimizing the separation distance) and is polarized perpendicular to the surface. The enhancement field drops off as r^{-3} [11], i.e. the signal intensity decreases as r^{-6}, where r is the distance of the molecule from the metal surface.

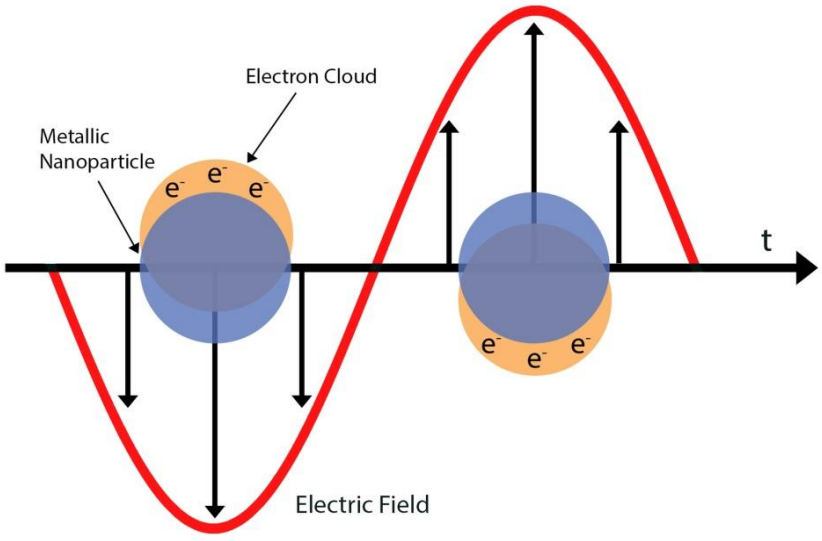

Figure 1.2 Schematic of the plasmonic enhancement effect in metal nanoparticles where the free electrons experience oscillations when excited by an electric field. The oscillations cause an increase of the electric field immediately surrounding the metal surface.

1.3.2. Molecular Markers of Disease

Biomarkers are molecules in the body that can be found either in the blood, other bodily fluids, or tissues and are important molecular signatures of the cell phenotype that can guide diagnosis and treatment plans [12]. Biomarkers can come in different forms such as proteins (enzymes, receptors), nucleic acids, antibodies, and peptides or they can be a collection of changes in gene expression, proteomic, and metabolic signatures. They have multiple clinical uses including estimation of disease risk, screening for primary cancers, determining the type of tumor (e.g. benign or malignant), disease prognosis and disease monitoring [13].

Current methods of biomarker detection *in vivo*, including ultrasound, radiology, fluorescence endoscopy and magnetic resonance imaging (MRI), lack the sensitivity and/or specificity to determine tumor behavior [14]. The major challenge is to be able to classify the disease early. The current 'gold standard' for diagnosis is histopathological examination *ex vivo* by microscopy of a biopsy sample [15]. This method is limited as it involves tissue extraction and does not provide functional information about the cells.

Monoclonal antibodies have been developed for cancer detection by mainly radioactive means of either SPECT (single-photon emission computed tomography) or PET (positron emission tomography). For example, one that is in clinical use is Indium 111 capromab pendetide (ProstaScint), a monoclonal antibody for prostate-specific membrane antigen (PSA) labeled with a radioactive tracer [16]. It targets the overexpressed antigen that is present in many instances of prostate cancer and aids in the staging and detection of relapse through imaging.

1.3.2.1. Antibody structure

Antibodies, or immunoglobulins (Ig), are molecules produced by immune cells in the body that are used to identify invading pathogens for other immune cells to destroy. The antibody is designed in such a way that it has a high specificity for the antigen, or characteristic protein, on the pathogen's surface. Binding is allowed by the roughly 'Y' shape of the antibody: **Figure 1.3**. Focusing on IgG antibodies, which were used in this work, they are formed with two light chains

and two heavy chains, completing a complex that is about 150kDa. The amino terminus of the chains is located at the tips of the arms of 'Y' shape, or in the Fab or variable region, where the antigen-specific binding sites are found. The carboxyl terminus is located on the base of the 'Y' which contains the Fc or constant region of the antibody. In immunohistochemistry, this Fc region is recognized by a secondary fluorescently-tagged antibody or Fab fragment.

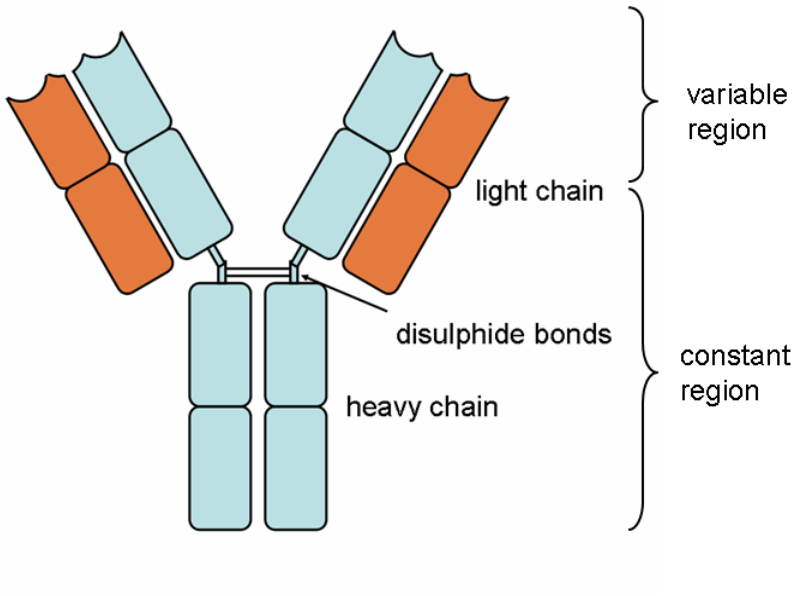

Figure 1.3 Schematic of an IgG molecule containing 2 heavy chains and 2 light chains making up the constant and variable regions.

1.3.2.2. Lung Cancer as a Model

Lung cancer has a very dismal 5-year survival rate of 20% [1] and most of the patients are at a late stage of the disease upon diagnosis. For patients diagnosed at Stage I the 5-year survival rate is higher at over 60%, but this represents only a small proportion of patients. The challenge in this disease is early detection – the patients need to be diagnosed before the disease grows and metastasizes.

A recent National Lung Screening Trial [14] showed that screening using low-dose helical computed tomography decreased mortality in patients with a high-risk of developing lung cancer by 20% compared to screening using chest radiography. Low-dose CT had a rate for positive screening test of 24.2% compared to 6.9% with radiography. Both of the tests had a high false-positive rate of 96.4% (low-dose CT) and 94.5% (radiography). This leads to abnormalities that need to be further investigated. This can be done by further imaging, bronchoscopy, biopsy or surgical procedures. A novel way to characterize these abnormal lesions is to use molecular markers.

Lung cancer can be divided into small cell lung cancer and non-small cell lung cancer (NSCLC), the latter being more common with 85% of cases being this type [1]. NSCLCs are known to be a group of distinct diseases with genetic and cellular characteristics that are different [17]. Thus, for the identification of lung cancer lesions a spectrum of molecular markers must be used for correct classification.

Lung cancer has been found to have mutations in different types of proteins and receptors. HER2 is member of the EGFR family of receptor tyrosine kinases and HER2 mutations are found in 2% of NSCLC [18]. MET amplification was detected in 22% of lung cancer specimens that developed resistance to gefitinib or erlotinib, which are drugs developed to target tumors with EGFR-activating mutations [19]. PIK3CA is a common gene overexpressed in cancer cells. It was found to be overexpressed in 33% of squamous cell carcinoma, 6% adenocarcinoma and 5% of small cell lung cancer cell lines [20].

1.3.2.3. Quantum Dots and Fluorescent Markers

An established method to determine the expression levels of antigens in cells is to use a fluorophore conjugated to a targeting antibody. A fluorophore is a small molecule that emits a photon with a longer wavelength than a photon by which it was initially excited. The fluorophore is conjugated to the antibody which then binds to the sample. The collected fluorescence signal represents the amount of bound antibody indicating the presence of the antigens in the sample.

The intensity of the fluorescence signal reveals whether there is high or low expression of the associated protein.

Conventional fluorophores are bright, small molecules that are well established and easy to use. The downside is that they easily photobleach after prolonged exposure to an excitation source and they have a broad emission spectrum making it difficult for spectral separation of different types in a combined sample. This does not allow for quantitative analysis of the antigen expression levels, due to the changing signal intensity and the inability to accurately determine the amount of signal originating from each type of fluorophore in the collected spectrum.

Quantum dots are an improvement on standard dye-based fluorophores. They are crystals of semiconductor material that have a tunable emission wavelength that varies depending on the size of the quantum dot. Their signal is stable and they have narrow emission spectra, allowing for a more quantifiable way of separating the spectra from the different types of quantum dots. Bioconjugated quantum dots have been used for quantitative immunohistochemical analysis [21,22]. The potential *in vivo* applications are limited, since quantum dots are typically composed of heavy metals such as Cd and Se, and so are potentially toxic [23].

It is preferred that the optical marker would be active in the 690-900 nm range, which is the ‘tissue optical window’. At this range light penetration in tissue is maximized. Hemoglobin in the blood absorbs strongly below about 600 nm, which limits the signal intensity, due to decreased penetration of the excitation and emission light. Water absorption is also low in the near-infrared range [24]. However, there is a limited number of fluorophores available in this range which limits their use in detecting multiple biomarkers simultaneously.

Thus, there is a need to develop an optical probe to be usable *in vivo* in the ‘tissue optical window’ and that can be easily and quantitatively unmixed spectrally to be able to determine the presence of multiple antigens simultaneously.

1.3.3. SERS NPs as optical markers

SERS has been exploited to create nanoparticles (NPs) that can be used as optical markers that have an associated Raman spectrum. Due to their strong signal they can be detected on tissue, they possess low toxicity and have the ability for multiplexed detection. SERS NPs are seen as a potential *in vivo* multiplex molecular imaging probe. This section will describe the use of SERS NPs as optical probes in medicine and the progress that has been made to date towards them becoming a clinical tool.

1.3.3.1. SERS NP Substrates

SERS NPs used for molecular targeting are typically synthesized to have a metal core, either gold or silver, and coated with dye molecules or small Raman -active molecules. Usually, a reference molecule is chosen that is small and does not have too many bonds, as increased bonding gives an increased number of Raman peaks in the spectrum, making the signal more complex, **Figure 1.4**. In terms of multiplexing, this causes a greater amount of peak overlap and increases the error in spectral component separation. The metal core acts as a Raman signal amplifier through surface plasmon resonance enhancement of the electric field.

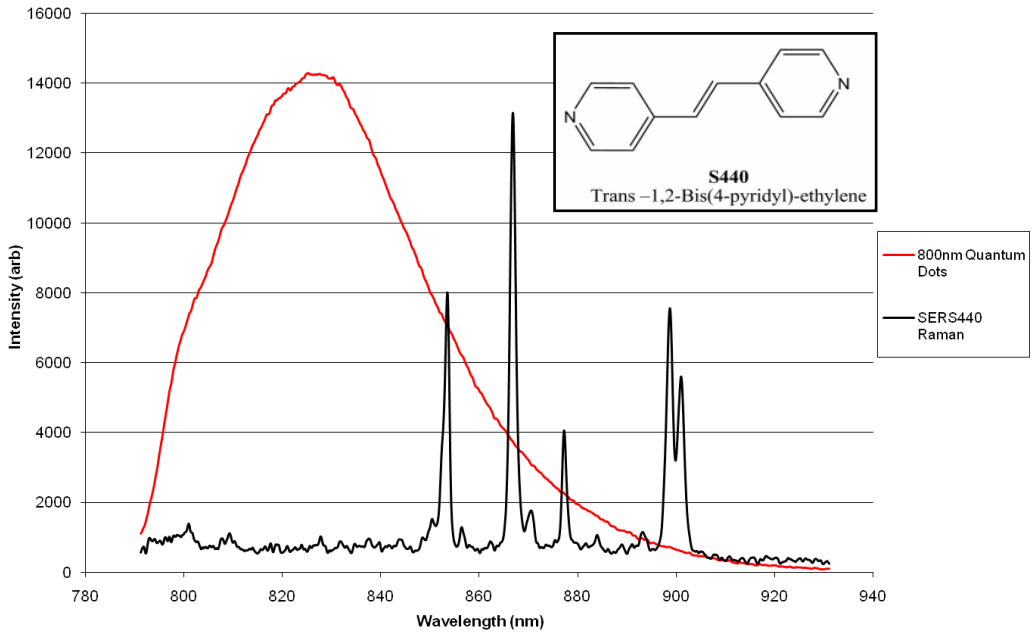

Figure 1.4 The Raman scattering spectrum of an S440 reporter molecule (trans-1,2-Bis(4-pyridyl)-ethylene) compared to the fluorescence spectrum of 800 nm quantum dot. The Raman peaks are much narrow, making them easier to separate spectrally. (Courtesy of Patrick Z. McVeigh)

The next layer acts as a stabilizer of the NP. It is commonly composed of either polyethylene glycol (PEG) molecules or a silica shell. The silica shell provides greater chemical stability of the Raman reporter molecules that are adsorbed to the core, especially when encountering a range of environments for *in vivo* applications [25]. The PEG chains, on the other hand, are more permeable to encountered molecules and allow them to approach the metal core. This can change the SERS reference spectrum, as the Raman spectrum of the encountered molecule will also be enhanced.

The next layer is added to reduce the nonspecific binding that occurs mainly due to the charges of the Au NP. Strategies to reduce nonspecific binding have shown that PEG is superior for *in vivo* applications. PEG has a long history of being used to prevent absorption of biological

molecules and prevent uptake by the reticuloendothelial or macrophage system [26]. Uptake in other tissue decreases the NPs that are able to reach the tumor site. Therefore, a PEG coating of surrounding the silica shell of the NPs is necessary to limit the electrostatic interactions. A further advantage, is that the PEG molecules can be formulated with reactive groups at the ends that allow for functionalization of the NPs with antibodies or other moieties [27] for targeted delivery. This facilitates the binding of functionalized SERS NPs to the cell surface that leads to expression level characterization.

1.3.3.2. Multiplexing testing

With increasing proteomic information becoming available, it would be advantageous to use this for diagnosis of disease *in situ*. Cell surface markers, or antigens, can have a change in their expression levels indicative of transformation into malignant cells. To be able to determine which types of antigens are present, and in what proportion, will allow the patient to be diagnosed more precisely, leading to a more personalized treatment directed at the specific disease subtype. In lung cancer, for example, EGFR can become overexpressed or mutated in 30% of squamous cell carcinomas but only 15% of adenocarcinomas [28,29]. Other proteins can also become amplified in other types of lung cancer, as discussed previously. Multiplexing is advantageous as it allows for the simultaneous detection of a range of biomarkers. In terms of targeted SERS, quantitative multiplexing is the determination of the proportion of bound SERS NPs that are targeted to a specific biomarker.

SERS multiplexing without targeting has been reported by multiple laboratories. When combining different types of SERS NPs in a mixture, the collected Raman spectra can be accurately unmixed, that is, spectrally separated, to reveal the proportion of each of the SERS NPs present. The relationship between the concentration of the SERS NPs and the signal intensity from a Raman reference molecule is linear, allowing direct correlation of the signal to the amount of NPs present. This is under the assumption that the NPs are stable, i.e. that the number of Raman reporter molecules adhered to the Au core is constant and there is no aggregation of the NPs. A change in these factors will change the intensity of the Raman signal and conclusions cannot then be drawn as to the amount of NP present [30]. A step forward from this, is showing that multiplexing is valid when collecting the signal from SERS NPs located on

biological tissue, which can contribute a strong autofluorescence background signal. This was shown in our lab by McVeigh *et al*. [30] using narrow-band SERS imaging, by applying a mixture of four different types of SERS NPs through a subcutaneous injection on the dorsum of a nude mouse: **Figure 1.5**. A quadruplex mixture of varying concentration was successfully spectrally unmixed.

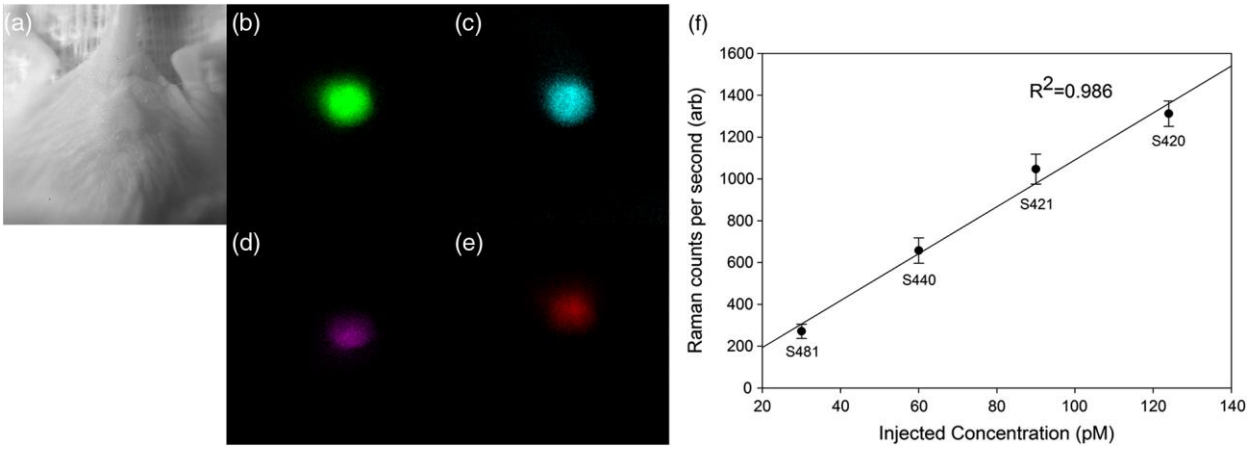

Figure 1.5 White light (a) and bandpass images (b-e) of the 4 types of SERS NPs in varying concentrations (S420, S421, S440, S481) applied through a subcutaneous injection on the dorsum of a nude mouse. Unmixed Raman bandpass image intensities as a function of known intensities in a quadruplex mixture are shown in (f) with the error bars representing ±1 s.d. (Reprinted with permission from ref. [30], Copyright 2009 Society of Photo Optical Instrumentation Engineers)

Further studies have shown that targeted SERS NPs can bind to cancer tissue following either intra-tumoral injection or topical application. Dinish et al. [31] showed that intra-tumorally injected SERS NPs targeted to three different cell surface antigens were able to bind to a subcutaneous breast cancer xenograft tumor. The analysis was done only semi-quantitatively, where the presence or lack of binding was determined just by the intensity of the Raman peak at the specific wavenumber. Wang et al. [32] were able to show that quantitative triplexing can be achieved using two targeted SERS NPs and one untargeted to account for non-specific binding.

This was done using a fiber optic probe and a brief topical application. In our lab, Mallia *et al.* [33] showed, using narrow band imaging, that it is possible to spectrally unmix a triplex of SERS NPs in which 2 were targeted with 2 different anti-EGFR antibodies, as well as an untargeted control. This work showed that targeted binding does occur by having very similar binding with the two anti-EGFR probes and no binding with the unspecific probe, **Figure 1.6**. Another conclusion was that the type of SERS NP reporter molecules used does not affect the antibody targeting or the collected signal.

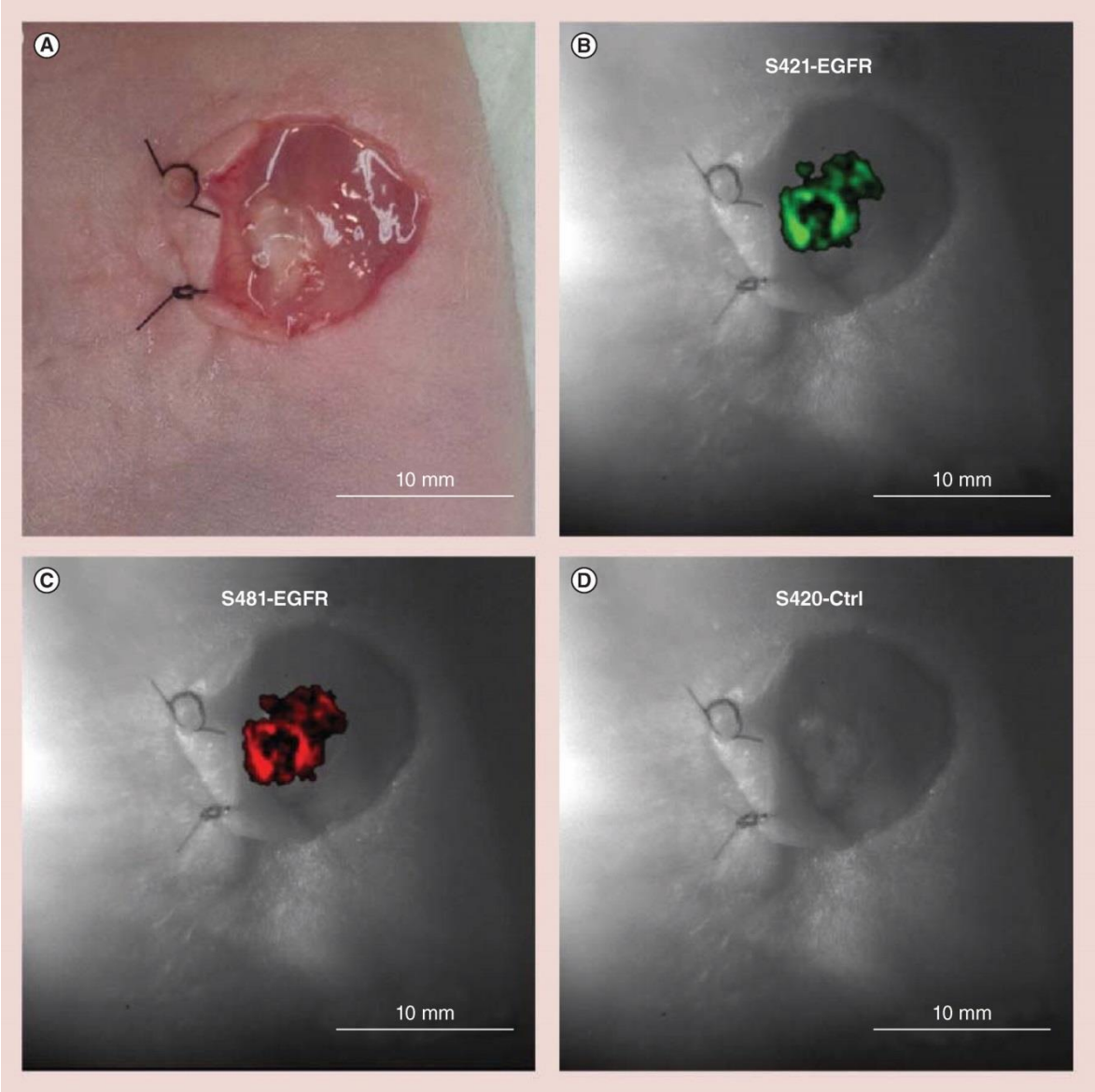

Figure 1.6 *In vivo* tumor targeting and multiplex detection of the EGF receptor using wide-field Raman imaging in a lung cancer (A549) xenograft tumor. (A) Color image showing topical application of the nanoprobe cocktail on the surgically exposed tumor. Wide-field Raman bandpass images shown in false color superimposed on the white light image correspond to (B) cetuximab labeling of S421 nanoprobes, (C) panitumumab labeling of S481 nanoprobes and (D) nonspecific labeling of control S420 nanoprobes, where S421, S481 and S420 SERS NPs are associated with three different Raman reporter molecules. (Reprinted with permission from Ref. [33], Copyright 2014 Future Medicine Ltd)

1.3.3.3. Endoscope development

SERS NPs can be employed to enhance the signal and increase the contrast for normal and diseased tissue. To detect optical markers *in vivo*, endoscopic techniques have been in development for standard white-light imaging. To be able to collect the signal for specialized optical probes, different light collection methods need to be adapted. This was discussed in a book chapter entitled "Nanoparticle-Enabled Optical Endoscopy: Extending the Frontiers of Diagnosis and Treatment" authored by B.C. Wilson and S. Borel [34]. This section will focus on imaging technologies developed for collecting SERS spectra.

The use of targeted SERS NPs is tending towards topical application in body cavities using endoscopic imaging. This minimizes the dose that needs to be given to the patient and greatly reduces systemic uptake of the nanoparticles. Specific instrumentation needs to be developed to detect the SERS signal in a clinical setting, with fast acquisition time and image formation. Currently employed standard white-light endoscopes in the clinic have poor performance in the infrared wavelength range, so that SERS detection requires a modification in the detection cameras and optics.

Currently there are two main signal collection methods, full spectrum collection and narrow-band detection, each with their own advantages and limitations. Full-spectrum collection takes longer but provides more detailed information. Direct classical least-squares (DCLS) fitting or principal component analysis (PCA) is applied to the measured spectra to determine the

components and the relative amount of bound SERS nanoparticles. This method is able to accommodate a wide range of background spectra that would be collected from the tissue autofluorescence, enabling more accurate SERS NP quantification [35,32,36].

Narrow-band detection collects only a specific narrow range of wavelengths that are representative of the SERS NP references and the background. This can be challenging, due to the highly fluctuating background spectra. The advantage is that this method is potentially faster since point spectroscopy (where spectra are collected at single locations) is avoided. An image is formed by collecting the light of a particular intensity using a CCD (charge-coupled device). The image is collected for a wavelength range specific to the reference spectra (usually for a strong intensity peak), using a narrow-band filter, and the same is done for a wavelength range representing the background. After finding the difference of the two images, an image is formed representing just the reporter intensity without background interference. When multiple reporters are used, it is important to have reference spectra where the strong peaks for each are located far from each other, in terms of wavelength. This allows for more accurate subtraction with lower interference from other components [30]. Thus, this technique is limited by the availability of Raman reporters with a low number of Raman peaks.

1.4. Intrinsic Raman

1.4.1. Intrinsic Raman Spectroscopy

As a vibrational spectroscopic technique, Raman spectroscopy is an emerging biomedical tool that has the potential to provide a wide array of information about the chemical composition, molecular structure and molecular interactions of cells and tissues. It can be employed in a non-invasive, non-destructive and non-contact way for *in vivo* evaluation [37]. In biological specimens, the Raman spectra reflect the constituents of the tissue such as proteins, nucleic acids, lipids, carbohydrates and inorganic crystals. Since disease of normal cells and tissue leads to changes in cellular function and metabolic characteristics, it is possible to detect differences in the Raman spectra that are indicative of the disease. If such changes are unique to the disease type then they can be used as a phenotypic marker for characterization. Raman spectroscopic

evaluation can be done under normal physiological conditions and quickly, with minimal sample volumes.

Many research groups have chosen to use this information in classifying cells or tissue for a wide range of problems, for example, the effects of ionizing radiation exposure on tumor cells [38], diagnosing gastrointestinal lesions [39,40] determining malignant tissue margins in the brain [41] and categorizing at the single-cell level whether the cell is malignant or normal [42-44] and identifying stem cells [45,46].

Raman spectra in diseased tissue rarely differ by substantial changes in the peaks, with the exception of tissue calcification where the peaks are quite prominent. Nonetheless, the changes observed are significant and have been shown in many situations to agree with the gold standard of histopathology. **Figure 1.7** shows an example spectrum of breast cancer cells and the peak assignments for common constituents. The changes typically seen are subtle and involve small peak intensity changes in different wavenumbers, mainly involving protein, DNA, or lipid peaks. The changes can be expressed as changes in ratio between the peaks of different biochemical components.

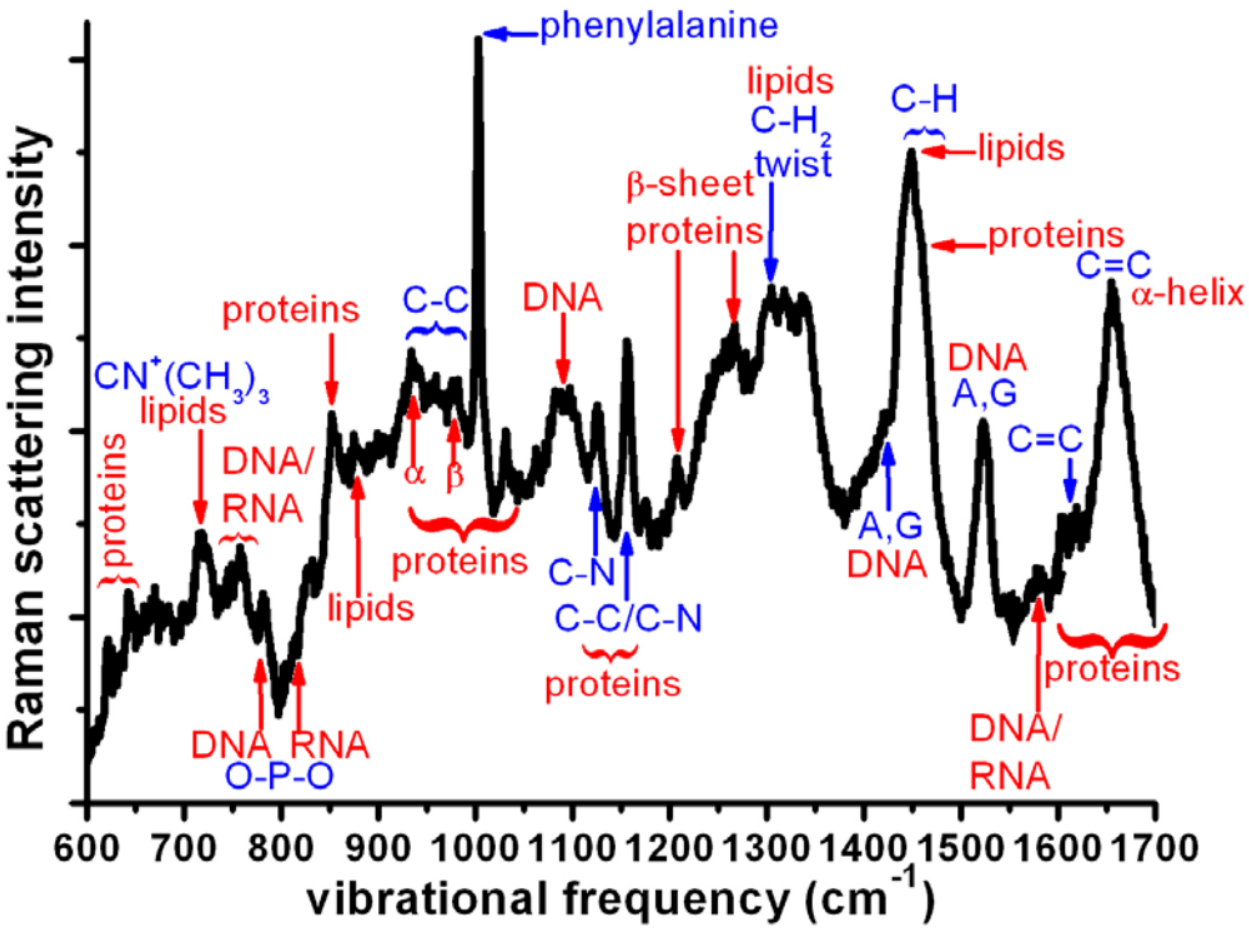

Figure 1.7 Raman spectrum of live MCF-7 breast cancer cells with peak labels representing Raman active vibrational levels in the bonds of proteins, DNA/RNA, and lipids. The vibrational frequency (or wavenumber) refers to the shift in frequency of the Raman scattered light relative to the excitation light, due to the molecular vibrational energy levels. (Reprinted with permission from ref. [47], Copyright 2010 A. Downes and A. Elfick)

Recent successes in using Raman scattering in biomedical applications has relied on the development of multivariate statistical analysis of the complex spectra. The Raman peaks from different molecular vibrations overlap, making it difficult to classify cells simply by evaluation of individual peak intensity ratios and locations. Multivariate statistical analysis allows for a much more detailed analysis of the spectra and more robust classification. Some methods that are commonly used are: least-squares fitting, where the Raman spectra of known chemical constituents are used to fit the data; principal component analysis (PCA), where the spectra are

transformed to a different orthogonal space to reduce correlation and redundant data; cluster analysis, where data clusters are established based on spectral similarity; and blind-source separation that is an unsupervised method for determining pure components in a data set [48].

1.4.2. *In Vitro* Raman Microspectroscopy

The ability to analyze cells on the single cell level is becoming increasingly important for our understanding of cellular biology. It is thought that cancer begins with a mutation in a single cell [49] that grows into a tumor and ultimately spreads to other parts of the body. Our understanding of the processes that enable this growth, as well as detection of these cells, is crucial to fighting the disease. Single-cell analysis is routinely done using optical microscopy. It has high resolution and sensitivity and is non-invasive. The disadvantage is that it needs to be combined with fluorescent or other labeling to be able to classify the cells. Electron microscopy provides even higher spatial resolution but requires complex sample preparation that cannot be done for live-cell imaging [50]. There is also a need for contrast agents to be able to identify specific molecules in cells, which depends on finding markers with high specificity for the characteristics of interest. Raman microspectroscopy on the other hand is a label-free technique that can detect intrinsic biochemical differences in the cells.

To date many types of cellular changes have been classified using Raman spectroscopy combined with multivariate statistical analysis. Cell-cycle variations have been successfully studied using chemometrics and linear-discriminant analysis (LDA) [51]. It has also been shown that the principal components in PC analysis of cell spectra capture the phase-to-phase transition differences of the cell cycle as ratios of protein and nucleic acid content to lipid [52]. Raman spectra have been used to determine drug effects on cells, for example the response of lung adenocarcinoma cells to the chemotherapeutic drug cisplatin [53], the effects of non-cytotoxic doses of the anti-cancer drug gemcitabine [54], and doxorubicin response of T-lymphocytes [55]. Another important application is the identification and classification of single cells. Raman spectroscopy has been shown to distinguish tumorigenic and normal cells for many types of cell lines where this method could be used for label-free cell sorting [45,56,57].

1.4.3. Applications for Disease Detection

The use of Raman microspectroscopy for *in vivo* work is technically challenging, as the Raman signal intensity is generally weak. Only about 1 in 10^8 incident photons are inelastically scattered compared to the majority being elastically scattered. Tissue autofluorescence, even in the near-infrared (NIR) can also be significantly stronger, and can overwhelm the Raman signal.

One of the first *in vivo* clinical applications of Raman spectroscopy was to investigate early cancer and dysplasia in the gastrointestinal tract [58]. Raman spectra of human gastrointestinal tissues were measured during routine clinical endoscopy, using a fiber-optic probe that was inserted into the instrument channel of the endoscope and placed in gentle contact with the mucosal surface of the tissue. This initial exploration revealed some differences between normal and diseased tissues but the data sets were too limited to draw definitive conclusions. Further work in this area showed that adenomatous and hyperplastic polyps in the colon could be distinguished with 95% accuracy [39].

As a further example, complete surgical resection of malignant brain tumors (gliomas) is extremely challenging due to the difficulty of delineating tumor tissue from normal tissue in the resection bed. Tissue preservation is very important for normal brain function, so that the extent of resection needs to be minimized, while maximum tumor removal is critical to extend survival. Studies in murine models showed that the tissues can be distinguished using Raman spectroscopy by their chemical composition [59]. Initial clinical trials in humans have been done by our colleagues in Montreal and have been very successful: initial trials using a handheld Raman probe during surgery glioma tissue could be distinguished from normal brain tissue with a sensitivity and specificity of 93% and 91%, respectively [41].

1.4.4. Ovarian Cancer

The application of Raman spectroscopy in discrimination of oral, cervical, breast, brain and other cancers has been quite active. The exploration of ovarian tissue has been limited, with only a few

studies showing preliminary evidence of being able to distinguish normal and malignant tissue *ex vivo*.

Ovarian cancer is the fifth-leading cause of cancer death among women and the leading cause among gynecological cancers [60]. The 5-year survival rate for this cancer is only around 30%. In large part this is due to late stage diagnosis, where around two-thirds of patients are diagnosed when the disease has already progressed to stage III or IV and involves other organs in the peritoneal cavity [61]. If the disease is discovered and treated at stage I (confined to the ovary), the 5-year survival is as high as 90%. This drops down to 33% for diagnosis at stage III or IV. Current treatment involves an initial surgical debulking, followed by chemotherapy using agents such as platin compounds and taxanes [62]. Prognostic factors include the age of the patient, the stage of the cancer, histologic grade and the presence of residual disease after initial surgery.

1.4.5. SERS in biological applications

SERS provides characteristic spectral information from Raman scattering and has the advantage of a highly enhanced signal overcoming the problem of low sensitivity. Thus, it is able to provide 'fingerprint' information along with having a high sensitivity for signal collection. Biological samples, being composed of complex biomolecules, can be characterized using SERS. While another application is the use of plasmonic nanoparticle functionalization with Raman active reporter molecules and conjugation with targeted ligands (i.e. antibodies) for detection of specific molecules on the cells/tissue [3].

SERS can be utilized in aqueous solutions to detect certain components, with the specific advantage of the detection of low concentrations. This can be applied in detection/monitoring of bacteria [63], specific molecules [64,65], drug release [66] and cancer cells [67,68].

Gold nanoparticles can be mixed with cell media, which causes the signal collection time to be greatly reduced due to the enhancement effect. The enhancement of the electric field due to the oscillations of the surface plasmon in the nanoparticles amplifies the number of Raman scattered photons by orders of magnitude. This allows probing of the intracellular chemical composition if the gold nanoparticles are able to enter the cells. The intracellular concentration can be

increased by electroporation [69]. The SERS enhancement has been shown to be able to measure cell content in osteosarcoma cells [70,71]. **Figure 1.8** shows results for a blood serum sample from a colorectal cancer patient. The signal is greatly enhanced using SERS (**Figure 1.8a**) compared to the intrinsic signal with no enhancement (**Figure 1.8b**). Combined with PCA/LDA methods, the use of the enhanced spectra correctly differentiated normal and colorectal cancer cells in a blood serum mixture with a sensitivity 97.4% and specificity of 100% [72].

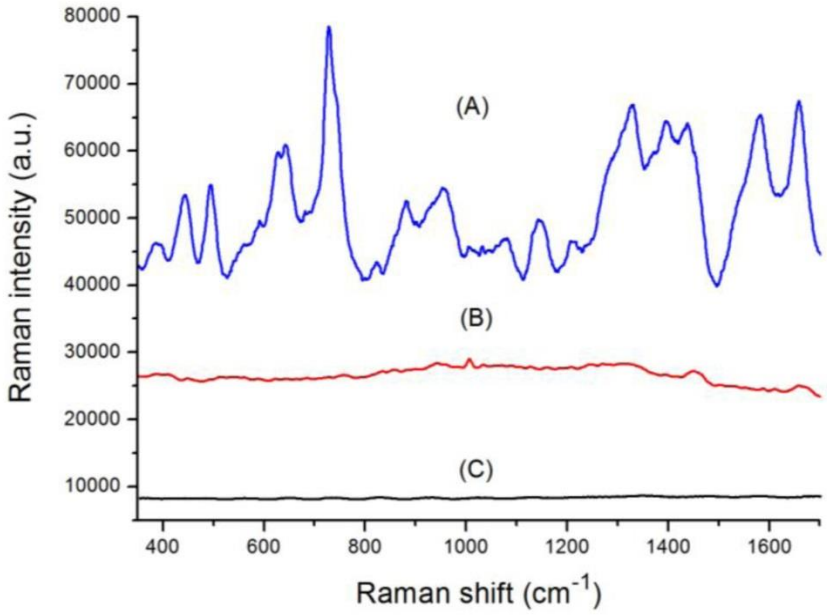

Figure 1.8 (A) SERS spectrum of the blood serum sample from a patient with colorectal cancer, (B) the regular Raman spectrum of the same serum sample without the addition of gold nanoparticles and (C) the background Raman signal of the coagulant agent mixed with gold nanoparticles. (Reprinted with permission from ref. [72], Copyright 2009 Society of Photo Optical Instrumentation Engineers)

1.5. Thesis Objectives

The overall goal of the work in this thesis is to contribute to the development of Raman spectroscopy as a diagnostic tool for cancer using both intrinsic Raman spectra of cells and targeted SERS multiplexing to molecularly diagnose the presence of cancer cells. The techniques

were tested *in vitro* in murine normal and transformed ovarian surface epithelial cells for intrinsic Raman, and in lung cancer cell lines using SERS multiplexing. These quantitative studies are intended for future translation into clinical use.

Chapter 2 describes the methods used to validate SERS NPs, followed by the testing of quantitative multiplexing of the targeted SERS NPs using lung cancer cell lines. **Chapter 3** of the thesis describes the development and testing of Raman microspectroscopy and multivariate statistical analysis in classifying live ovarian surface epithelial cells into two groups – normal and malignant. **Chapter 4** summarizes the results and describes the general trends in the field and the future of clinical Raman spectroscopy.

Chapter 2
Validation of SERS NPs

2.1. Introduction

Recently there has been a lot of interest in discovering more cell antigens that have diagnostic value. The search for better cancer treatment has really advanced the research and discovery of tumor specific antigens [73,74]. The same antibodies can be used for disease diagnosis. Research has shown that cancer can be very heterogeneous, so that an array of antibodies would work better to increase the accuracy of the diagnosis [75]. To be able to perform diagnostic tests *in situ*, probes may be conjugated to antibodies to detect their binding to associated antigens. In this case, fluorescence markers will not work, as they are difficult to multiplex quantitatively especially in the presence of background autofluorescence that is found in cells and, more significantly, tissue. SERS (surface enhanced Raman scattering) NPs provide much narrower peaks, by 1-2 orders of magnitude, and have the potential to multiplex a large number of antigens simultaneously due to this better spectral separation. Before this method can be translated to the clinic, it is important to determine its limitations. It must be confirmed that it is an accurate representation of the information on the cell surface.

Contrast agents based on (SERS) have been studied by several groups over recent years. The classes of agents include metal nanoparticles (NPs) with a Raman active probe adsorbed to the surface, the signal from which is enhanced through SERS. The Raman reporter molecule can be changed to obtain a different Raman spectrum creating different SERS NP types. The metal cores have been tried with many variations of materials and shapes, such as silver or gold nanospheres [76,77] or nanoshells [78] or even a combination of a gold and silver core and shell [79]. Various other shapes such as nanostars [80,81], nanorods [82,83], and others [84-86] have been investigated. The important characteristic for the metal core to work well for an optical probe is for it to provide a strong and consistent signal. Shapes that are irregular or have sharp edges create enhancement regions that are non-uniform around the whole surface making them not ideal for this application due to heterogeneous signal. They are also challenging to produce in consistent shapes for mass manufacturing that is necessary for adoption for clinical use. In terms of the material, gold NPs are preferred for *in vivo* applications due to their unreactive nature.

Commercially produced silica encapsulated gold SERS NPs have been used in this work. The silica encapsulation pacifies the NPs by increasing their chemical and physical stability. This also protects the Raman reporter molecules adhered to the gold core from interacting with the environment, as any change in their concentration will affect the SERS signal intensity. Silica encapsulation prevents aggregation of the gold nanoparticles or other molecules being adsorbed to the surface and altering the SERS signal. A reliable synthesis method for functionalization of these SERS NPs with monoclonal antibodies was recently developed [32] in our lab. The NPs were shown to be resistant to aggregation and signal degradation under different possible biological conditions. The conjugation process was shown to produce stable SERS NPs; therefore, this was not tested in this body of work.

These types of SERS NPs have also been reported by other research groups. Initial work included testing the signal strength and spectral separation of the signal in the presence of tissue background [87,88]. As this was successful, further work using SERS NPs conjugated to antibodies showed successful binding of the probes to antigen sites for one type [27,89,90] and several types of antibodies (conjugated to different SERS NPs) simultaneously [91-93]. These studies have not examined the limits of quantifiability of using antibody targeted multiplex SERS NPs. The aim of this work was to determine this, but in the process it was found that the antibodies used for conjugation to the SERS NPs need to be carefully evaluated for changes in binding affinity. This work found that some antibodies are not viable for use with this conjugation protocol and had greatly decreased affinity.

My contribution to the SERS NP research done in this lab is described in this chapter. By using a previously developed antibody conjugation method to SERS NPs, I was able to conjugate different types of antibodies to the NPs. I characterized SERS NP reference spectra after conjugation to antibodies and determined the relative number of conjugated antibodies between different types of SERS NPs. I developed and implemented a spectral unmixing method using least squares and tested its ability to accurately determine the concentration of SERS NPs present in the signal. I explored targeted SERS NP multiplexing for its ability to determine quantitatively the relative expression levels of 3 different cell-surface receptors, EGFR, HER2 and EpCAM, in 3 different human lung cancer cell lines that overexpress these receptors, namely MGH7, Calu-3 and H520, respectively.

2.2. SERS NP Conjugation

Commercially available SERS gold nanoparticles each with one of 4 unique associated Raman reference spectra were conjugated with a corresponding monoclonal IgG antibody and pacified with polyethylene glycol chains. The conjugation method is described in this section.

2.2.1. Materials

SERS nanoparticles were purchased from Cabot Security Materials Inc. (Mountain View, CA, USA). They have a 60 nm gold core, covered by adsorbed Raman active molecules, and a 30nm thick surrounding silica shell. The shell contains a thiolated layer that makes the NPs easy to functionalize with targeting molecules. The 4 different types of adsorbed Raman active molecules contributing to the detected spectra are S420 (4,4'-dipyridyl), S421 (d8-4,4'-dipyridyl), S440 (trans-1,2-Bis(4-pyridyl)-ethylene), and S481 (4-Azobis(pyridine)).

The polyethylene glycol (PEG) crosslinker, MAL-PEG_{24}-NHS ester, with a reactive maleimide (MAL) group and an N-hydroxysuccinimide (NHS) reactive ester group was purchased from Quanta BioDesign, Ltd. The PEG terminator chain, MAL-PEG_4-(m- PEG_{12}-)$_3$, was also purchased from Quanta BioDesign, Ltd. The anti-epithelial cell adhesion molecule (EpCAM, clone AUA1), anti- epidermal growth factor receptor (EGFR, clone H11), anti- human epidermal growth factor receptor 2 (HER2, clone 9G6.10) and an isotype control (clone MOPC-21) antibody were purchased from Fisher Scientific. The antibodies were all of the same sub-type of immunoglobulin – IgG1 – and are from a mouse host with human reactivity.

The reaction and storage buffers were prepared with (N-morpholino) propanesulfonic acid (MOPS), bovine serum albumin (BSA) and sodium azide. A solution containing 2-mercaptoethanesulfonic acid (MESA) was using to finish the reaction. All these supplies were purchased from Sigma-Aldrich.

2.2.2. Conjugation Reaction

A reaction buffer of 10mM MOPS, pH 7.25, was prepared. The antibodies were purchased pure in a PBS buffer and processed through a desalting column to remove any salts that will compete with the reaction. The reaction buffer, SERS NPs and antibodies, in a 1 NP to 350 IgG ratio were combined in a glass vial. The crosslinker MAL-PEG-NHS ester was added in a 50 PEG to 1 IgG ratio. The reaction proceeded for 3h in a table vortex at room temperature. The MAL end of the crosslinker reacted with the thiol groups on the NP surface and the NHS ester end reacted with the amine groups on the antibodies. Upon completion of the first reaction, PEG terminator chains were added in a ratio of 650,000 PEG to 1 NP. The reaction proceeded overnight at 4℃. It was quenched using 2-mercaptoethanesulfonic acid (MESA) to remove any unreacted thiol groups on the NP surface. To purify the suspension of any unreacted molecules, rinsing by centrifugation was performed 4 times. The product was stored at 4℃ at a concentration of ~3000 pM in a 20 mM MOPS buffer (pH 7.5) with 0.5% BSA and 0.05% sodium azide.

2.2.3. Conjugation Products

The final product was the conjugation of the 4 different types of SERS NPs, each having its own Raman spectrum from the Raman reporter molecule, with the corresponding antibody. They are as follows: S420-(anti-EpCAM), S421-(anti-EGFR), S440-(isotype control) and S481- (anti-HER2).

2.3. Instrumentation

For Raman signal collection a Renishaw inVia Raman spectrometer connected to an inverted Leica microscope was used. All the Raman spectra were collected using this system with the microscope objectives or the cuvette solution measurement system, as described in the following sections.

A 785nm incident laser beam was used to excite the SERS NPs. The signal was collected within a narrow spectral window of wavenumber shifts ranging from 600 to 1700 cm^{-1} (corresponding to 824-906 nm). Depending on the sample, the acquisition time was varied to achieve good signal-to-noise. For solution measurements this was 10s times 5 accumulations and for area measurements on adhered cells it was 64 s for the whole sample area (0.2 s per pixel).

The initial setup of the Raman signal collection system contained a microscope stage. This is optimal setup for measuring samples of adhered cells on slides or plates. For SERS NP or cells bound with targeted SERS NP solution measurements, the signal can be collected more easily using a cuvette system. The setup for this system was designed and installed as an addition to the Renishaw system.

A cuvette holder was purchased from Thorlabs Inc. The holder was installed on the Renishaw system in the collimated laser beam path from the spectrometer and before the beam reaches the microscope. The holder contained openings for the light path, a slot for a cuvette and a slot to hold a filter. The filter holder was repurposed to hold a plano convex lens (diameter 6.0mm, focal length 10.0mm) that focuses the beam in the sample. The focused laser spot was located between the middle and the further side of the cuvette avoiding any signal collection from the cuvette material itself. This allowed for the use of plastic cuvettes without interference from the plastic Raman signal. The backscattered Raman signal from the sample was collimated back through the lens and collected by the spectrometer. When cuvette measurements were not required the lens was taken out and the laser path was clear to be used in the microscope.

2.4. Validation of Conjugation

To be able to use the conjugated SERS NPs *in vitro*, they were first tested for consistency by obtaining the final concentration after conjugation and the relative amount of antibodies bound for each type of NP.

2.4.1. Materials

A secondary antibody marker of Alexa Fluor 647 fluorescent dye conjugated to anti-mouse IgG (H+L) $F(ab')_2$ fragment (Cell Signaling Technology) was used to determine primary antibody conjugation to the Au SERS NPs.

2.4.2. Methods

2.4.2.1. Equimolar Concentrations

To confirm a stable conjugation of the antibodies to the SERS NPs without aggregation, the absorption spectrum of the product was taken. If the Au NPs have aggregated, then the absorption spectrum will shift from the expected spectrum of the 60nm Au NP. The absorption peak of the unconjugated SERS Au NPs is 545 nm, which is characteristic of 60 nm diameter Au NPs [94]. The peak of the absorption spectrum was found to be 547 nm, which is slightly red shifted. This may be due to the addition of PEG molecules surrounding the nanoparticles that makes them larger. The shift of the peak to even higher wavelengths would be indicative of aggregation and can also be observed in the solution by eye by the formation of aggregates in the suspension resulting in a faster setting time.

For analysis of unknown concentrations, the Raman reference spectra intensities of the conjugated SERS NPs at known equimolar concentrations are needed. The conjugated SERS NPs were diluted to 10pM in 0.5% BSA in PBS. Since there is still pipetting error associated with creating these solutions, the absorption spectrum peak intensity at 547nm was used for internal consistency as a better representation of the concentrations of each of the solutions relative to the others. The Raman spectra of the same solutions were collected using the cuvette holder on the Renishaw system. The spectra were collected over 5s x 4 accumulations with 3 repetitions after mixing the suspended particles with a pipette. The Raman spectra were normalized to the intensity values acquired from the absorption spectra. This process gives

accurate Raman reference spectra acquired at identical concentrations that can be used during evaluation of all other Raman spectra.

2.4.2.2. Relative Amount of Antibody on SERS NPs

To be able to quantify the binding of the SERS NPs, it is important to know whether there is the same amount of antibody conjugated to each of the four types of SERS NPs. An indirect method with fluorescent secondary antibodies was used to determine the relative amount of antibody conjugated to the NPs,

Nine fmol aliquots of each of the SERS NPs were placed in 1.5mL Eppendorf tubes. 100µL of a 1:100 dilution of Alexa 647 conjugated secondary antibody was incubated with the NPs for 30 min in a table vortex for even binding and to prevent settling. The NPs were then rinsed 3 times by centrifugation and re-suspended in 0.5% BSA in PBS.

The fluorescence signal was measured using the Fluorolog Horiba system with excitation of the Alexa 647 fluorophore at 650 nm and emission signal collection from 660 to 750 nm. The absorption spectrum of the Au SERS NPs was collected from 360 to 800 nm. To determine the relative amount of fluorescence in the sample, the fluorescence emission peak intensity at 669 nm was taken to be directly proportional to the amount of bound fluorophore and, therefore, the amount of antibody conjugated to the surface.

2.4.3. Results and Discussion

The absorption spectra of the conjugated Au SERS NPs have peaks around 547 nm, **Figure 2.1**. Since the Raman reporter molecules do not have strong absorption peaks in this range and Au is a very strong scatterer, the signal can be attributed as coming from the Au NP core. The location of the peaks indicated a stable conjugation and no change in the absorption peaks were observed during storage for up to one month.

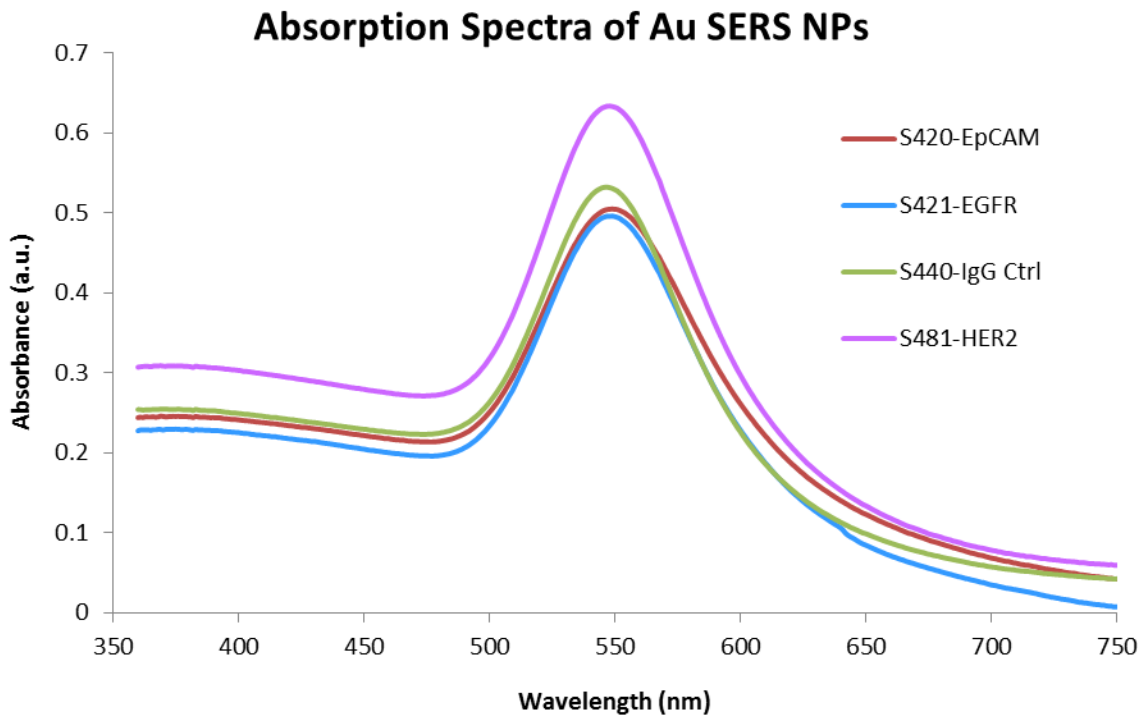

Figure 2.1 The absorption spectra of the conjugated Au SERS NPs.

The reference spectra of the four different types of SERS NPs were collected at equimolar concentrations: **Figure 2.2**. The S440 signal (green) is the strongest, so that for accurate results it was chosen to be conjugated to the isotype control antibody that was expected to have the lowest binding.

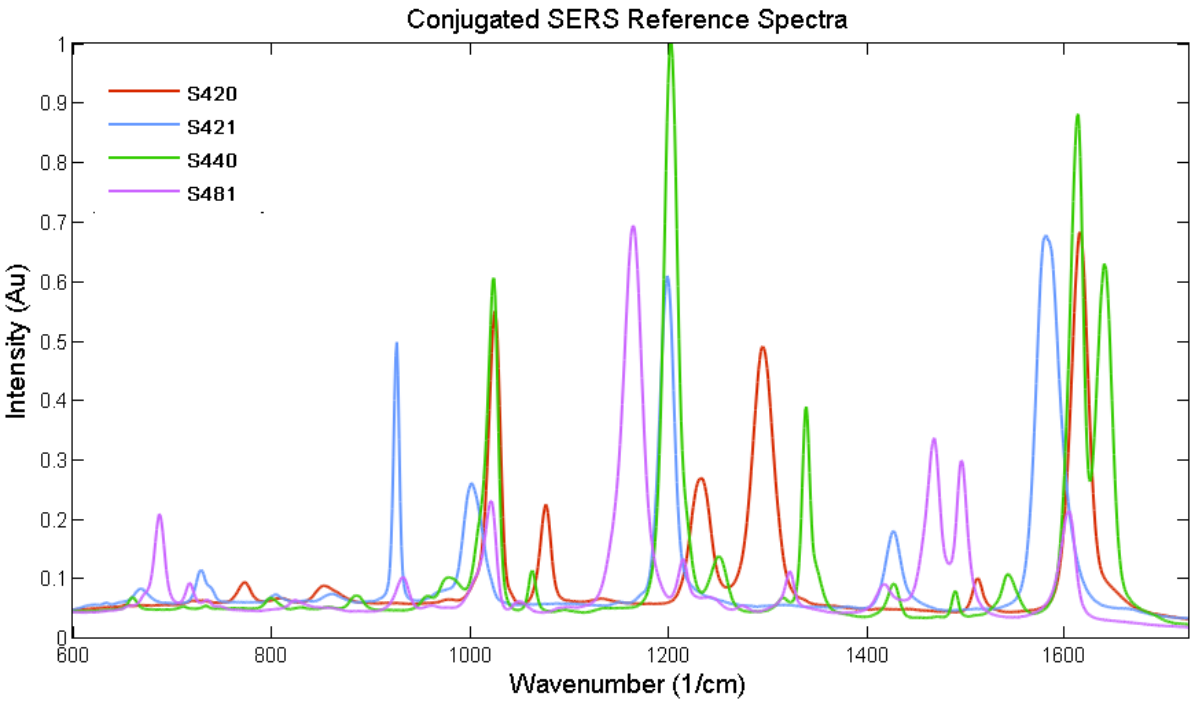

Figure 2.2 SERS NP spectra with normalized intensity (to the strongest peak of S440) at equimolar concentrations.

The normalized secondary antibody fluorescence in each of the four samples was found to be the same. **Figure 2.3** shows the fluorescence signal collected from the bound florescent secondary antibody to each of the NPs, indicating the relative amount of antibody conjugated. A control of SERS NPs with no conjugated antibody was used. These were conjugated with PEG terminator chains only, as for the antibody-conjugated NPs, which reduces non-specific binding to proteins and gives a similar structure of the control SERS NPs to the antibody-conjugated SERS NPs. The signal from the control sample was close to zero, as can be seen by the bar on the far right. The signal for antibody conjugated SERS NPs was much higher. This indicates that the secondary antibodies are binding specifically to the primary antibodies conjugated to the SERS NPs, and approximately the same amount of binding is seen for all four types of antibodies. Hence, it can be concluded that there is approximately the same amount of antibody conjugated to each type of SERS NP, i.e. that the antibody type does not affect the number of conjugates per NP.

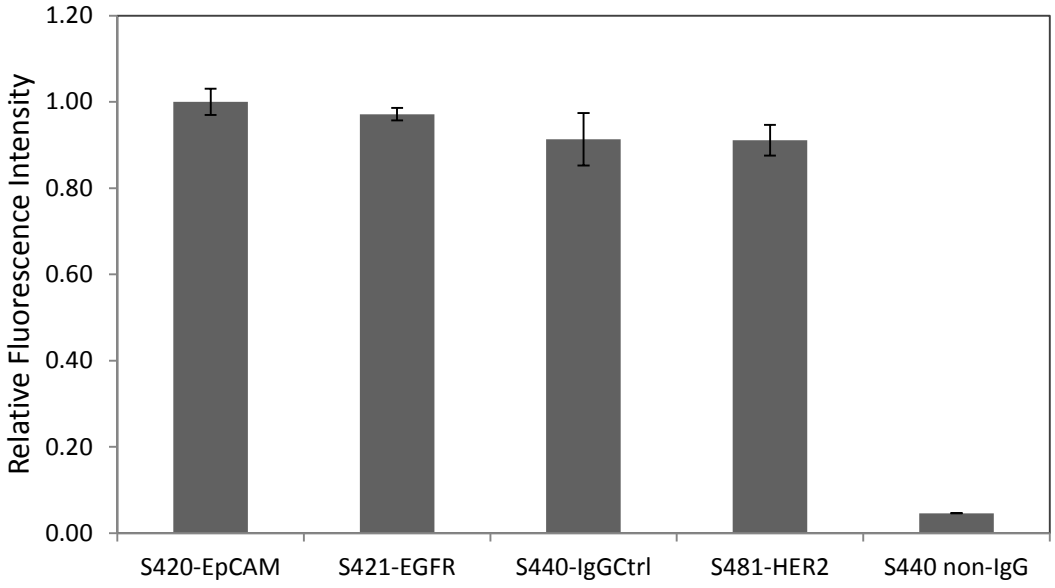

Figure 2.3 Relative fluorescence signal (normalized to the maximum signal intensity) from secondary antibody binding for each of the antibody conjugated SERS NPs (± 1 s.d., n=3).

The assumptions made when using equimolar NP concentrations for binding to the cells are that the NP concentration is proportional to the antibody concentration and is the same for each type of SERS NP. This assumption is assumed to be valid from the results in **Figure 2.3**.

2.5. Spectral Unmixing Validation

It is necessary to accurately determine the amount of signal present for each of the SERS NPs relative to the acquired reference spectra in order to quantify the amount of bound NPs on the cells. To validate the spectral component decomposition technique, Raman spectra of solutions with known NP concentrations were collected and analyzed using a least-squares algorithm.

2.5.1. Solution Preparation

Evaluating the accuracy of the least squares fitting requires careful allocation of the stock solutions to reduce pipetting error. All 4 types of antibody conjugated SERS NPs were diluted to

a 10pM concentration from the stock solutions. These solutions served to provide the reference spectra in the calculations. Volumes from these stock solutions were taken to make the mixed solutions of specific concentrations. The concentrations were chosen to reflect the expected results from cell staining. The S420, S421 and S481corresponding to EpCAM, EGFR and HER2, respectively, ranged from 10-60% of total SERS NP concentration in the solution and S440 corresponding to isotype control binding ranged from 2-25% of the total SERS NP concentration in the solution.

2.5.2. Measurement

The Raman signal of the solutions was measured in plastic cuvettes using the cuvette holder system installed on the Renishaw Raman system. The signal was collected for 10s x 3 accumulations, followed by re-suspension of the NPs using a pipette. The signal was averaged over 3 collections. This was repeated for all 10 solutions of different compositions.

2.5.3. Least-Squares Algorithm

Spectral decomposition was performed using the least-squares method algorithm in the MATLAB (MathWorks) software, which minimizes the difference, Q, between the model output values, $S(x_i, \boldsymbol{p})$, and the observed values. Here, *i* refers to the spectral points and p is the combination of fitting parameters, described as *a*, *b* and *c* below.

$$Q = \sum_{i=1}^{n} [y_i - S(x_i, \boldsymbol{p})]^2 \tag{2.1}$$

Apart from noise, the acquired spectrum is assumed to be composed of a linear weighted sum of the four different SERS NP spectra, a known background and a broadband baseline that is represented by a third order polynomial.

$$\boldsymbol{S} = \sum_{n=1}^{4} a_n \cdot \boldsymbol{F}_n + b\boldsymbol{B} + \sum_{\boldsymbol{n=0}}^{\boldsymbol{3}} c_n \cdot \boldsymbol{P_n} \tag{2.2}$$

$\boldsymbol{S}$ = measured spectrum

$\boldsymbol{F}_n$ = SERS reference spectrum for the n^{th} flavor acquired under the same conditions

a_n = weight for the n^{th} reference spectrum proportional to the concentration of SERS NPs

$b\boldsymbol{B}$ = weighted spectrum of known background, **B**

c_n = weight for the polynomial term

$\boldsymbol{P_n}$ = n^{th} order polynomial term

To evaluate the goodness of fit of the decomposed spectra, a spectral reliability index (SRI) was used as a metric [95]. A commonly used metric is the relative fitting error (RFE), which takes the norm of the fit over the norm of the acquired signal. A perfect fit will gave an RFE of 1 and a bad fit will approach zero. When fitting SERS spectra, the RFE can have a bias towards 1 if the background signal and the baseline signal are strong compared to the SERS signal. The SRI mitigates this bias by subtracting these two terms from the norm calculations.

$$SRI = \frac{\|S-R-(bB+\sum_n c_n \cdot P_n)\|}{\|S-(bB+\sum_n c_n \cdot P_n)\|} \tag{2.3}$$

The results presented below had a good fit based on $0.99 \leq \text{SRI} \leq 1$.

2.5.4. Results and Discussion

Comparing the expected fraction of SERS NPs present in the solution to the calculated amount using the least-squares fit of the reference spectra on the acquired signal, gives a result close to unity. A perfect result would be represented by a slope of one, m=1. Detection of the S420 and the S421 signals tends to be slightly higher than expected while for S440 it tends to be slightly lower than expected: **Figure 2.4**. The S481 signal is detected most accurately, with m=0.99. In general, the signal can be unmixed using the least squares algorithm with fluctuations of approximately ±5% from the expected values.

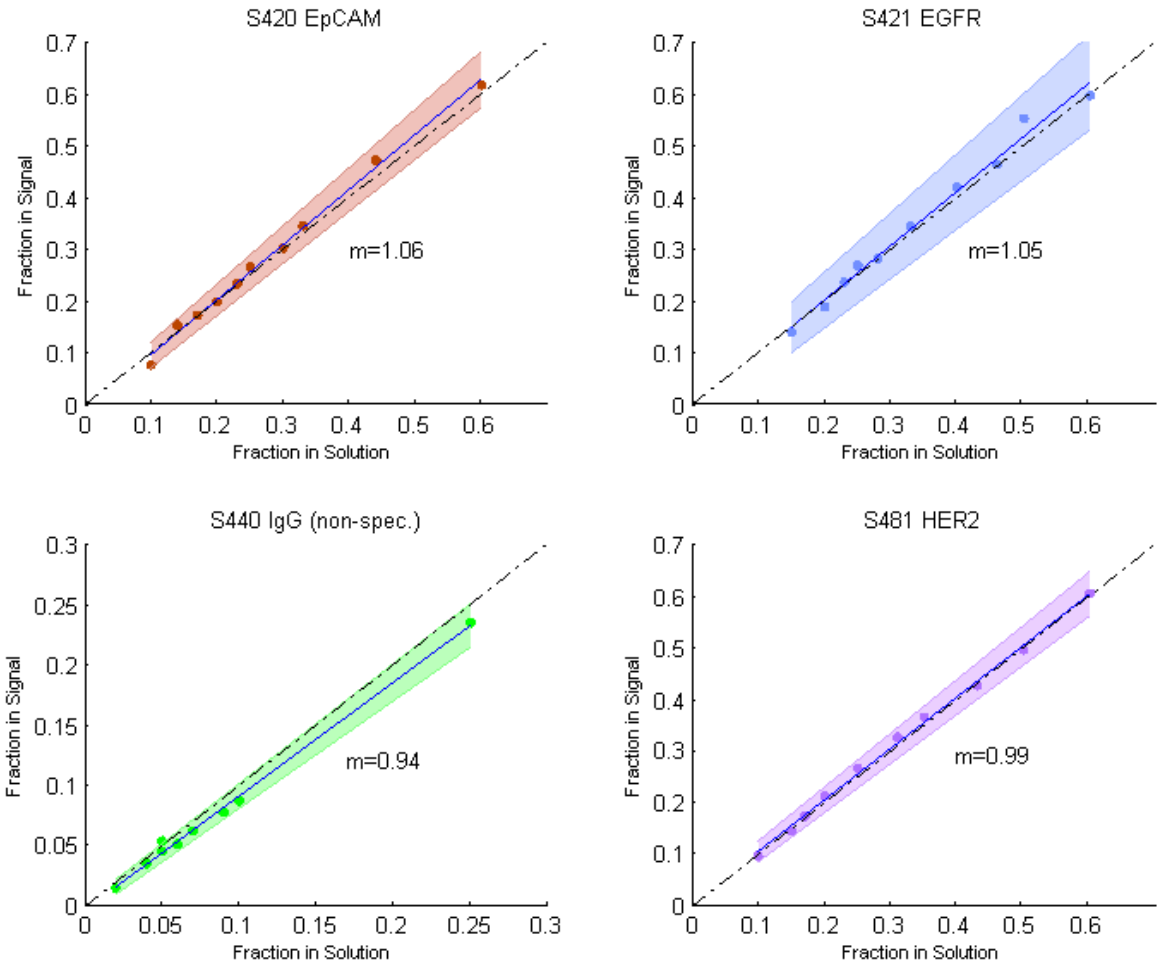

Figure 2.4 The calculated *vs.* expected concentrations of each of the four SERS NPs in a mixed suspension. The dots indicate the measured values with a line of best fit in colour. The dotted line represents x=y. The coloured area represents the 95% confidence interval.

The error seen when comparing the expected values to observed values may be due to the strength of the Raman signal from that particular Raman reporter molecule. For the same concentration of SERS NPs, the S481 and S440 signal intensities are the strongest: **Figure 2.4**, while the S420 and S421 signal intensities are weaker. Since the overall signal strength is different this can influence the uncertainty for that reporter. For S440, the concentration was kept lower than the other types, which made the signal intensities lower for this reporter in the solution. This gives lower signal for the algorithm to fit, which might tend to underestimate the

amount present. S481 signal is the most accurate, due to its overall signal strength and higher abundance in the solutions.

2.6. *In Vitro* - Materials and Methods

2.6.1. Cell Culture and Flow Cytometry

Calu-3, H520, A549 and MGH7 cell lines were cultured in RPMI-1640 media supplemented with 10% fetal bovine serum (Invitrogen, Burlington, ON, Canada). They were grown to confluency and detached enzymatically using Accutase (StemPro). Fixation was done using 4% paraformaldehyde in PBS for 15min. Fixed cells were stored in blocking buffer (3% bovine serum albumin (BSA) in PBS) at 4°C until further use.

Cells were prepared in the adhered state using 96 well plates. After detachment as described above, 10^4 cells diluted in media were added to each well. They were grown to confluency and fixed using 4% paraformaldehyde in PBS for 15 min and stored in blocking buffer at 4°C until further use.

Flow cytometry was performed to determine the expression levels of antigens on the cell surface. This was done using an indirect method. Fixed cells were aliquoted into assay tubes at 0.5 million cells per tube. Primary antibody incubation was done in 100 µL using a 1:1000 dilution of the 1mg/mL stock concentration of the monoclonal antibodies. The same monoclonal antibodies were used as for the SERS NP conjugation described previously. Following 30min incubation the cells were rinsed twice using a rinsing buffer (0.5% BSA in PBS). A fluorophore, Alexa 647, conjugated to a secondary antibody (1:1000 dilution) was incubated with the primary antibody-bound cells for 30 min. The same secondary antibody and fluorophore was used for all tubes. This was followed by 2 rinsing steps using the rinsing buffer. Flow cytometry was used to collect the fluorescence signal from each cell which indirectly represents the amount of bound primary antibodies. The mean of the signal for each event (total 10,000 events) was recorded and used as representative of the relative amount of cell surface antigens. A cell sample with no primary antibody was used as a control.

For comparing the antigen expression levels in their adhered state on a microplate, the same indirect staining protocol was used for cell preparation as for flow cytometry. First, cells were plated in 15 (5x3) wells in a 96-well plate. The first column of 3 wells was used as a background control and treated with only the fluorescently-tagged secondary antibody. The next 4 columns of 3 wells each were first treated with only one of the 4 primary untagged antibodies in a 1:1000 dilution for 30 min. This was followed by rinsing and treatment with a 1:1000 dilution of the fluorescently labeled secondary antibody.

The fluorescence signal was collected using the same Leica microscope as was used for the SERS signal collection on the Renishaw system. The filter corresponding to the 665nm Alexa 647 emission peak was used to collect the fluorescence signal. An image was taken of an area fully covered with cells with a cell density consistent for each cell line. The signal intensities correspond to the location and amount of the bound secondary antibodies. The intensity for the image was summed for each of the wells. This was done for all 4 columns of wells corresponding to the 4 different antibodies, with 3 repeats. A cell sample with no primary antibody was used as a control.

2.6.2. SERS in Microplates

The microplates with fixed cells were stained with a 40pM concentration of each of the types of SERS NPs in 3x6 wells. The first column of 3 was used as a baseline control for the background Raman signal. The next 4 columns of 3 were treated with 40pM concentration of each of the SERS NPs separately. The last column of 3 wells was treated with a mixture of all 4 SERS NPs, each at 40 pM concentration, to obtain the same number of each type of NPs as in the single wells . The represents the staining of the cells with all 4 markers simultaneously, which is the multiplex well, and this was compared to the other single wells for any interactions between the NPs.

2.6.3. SERS in Suspension

SERS NP staining was performed by aliquoting 0.5 million cells into Eppendorf tubes. The cells were re-suspended in 50μL SERS NPs at a concentration of 100 pM. They were placed for 30 min on a table vortex to prevent settling. To remove unbound NPs from the suspension, the cells were rinsed by centrifugation 6 times using the rinsing buffer. More rinsing steps were required when staining using targeted SERS NPs than with fluorescently labeled antibodies, due to the SERS NPs having a much higher mass. The final re-suspension was in the rinsing buffer and the suspension was then stored at 4°C.

The Raman spectra of the cells were collected on the Renishaw Raman system, for which the cell suspension was transferred to a plastic cuvette. The spectra were collected for 10s with 6 accumulations and this was repeated once after re-suspension. The spectra were then averaged.

2.7. *In Vitro* – Results Microplates

The first set of *in vitro* experiments was performed using fixed cells grown on microplates. This was chosen as the first step in trying to mimic the conformation of the cells in tissue, since the cells are adhered to a surface. The microplate set-up was chosen because of the accuracy in replicating the specific concentrations.

2.7.1. Fluorophore vs. SERS cell staining

Fluorescence microscopy was used to determine the antigen expression levels of the adhered cells. The fluorescence intensity of the antibody targeted cells compared to the non-targeted antibody treated cells, **Figure 2.5**, showed that the detected antigen levels were very similar to the expression levels obtained by flow cytometry of single-cell suspensions (shown further in section 2.8.1). This is the expected distribution of cells treated with SERS NPs: Calu-3 overexpresses HER2, H520 overexpresses EpCAM and A549 overexpresses EGFR.

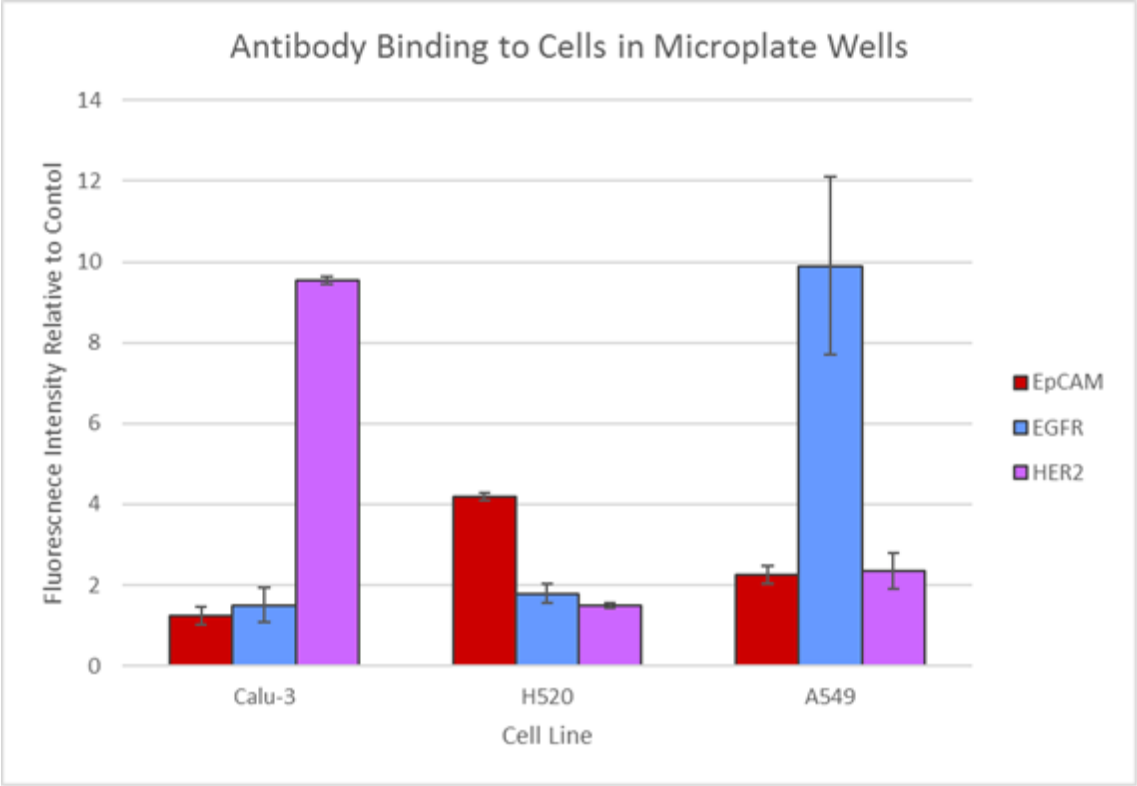

Figure 2.5 Targeted antibody (EpCAM, EGFR, HER2) binding relative to non-targeted antibody binding in Calu-3, H520 and A549 cell lines. (N=3 repeats, with bars indicating ±1 standard deviation).

Comparison of the signal collected from the fluorescently-stained cells to the SERS NP binding measurements was not as expected. The relative signal distributions corresponding to the different targeted antibodies did not exactly parallel the free-antibody binding. Representative results from the Calu-3 cell line can be seen in **Figure 2.6**. For this cell line, the individual wells where cells were treated with only one type of SERS NP had a similar distribution to that expected. In the multiplexed well, QUAD, the results showed a slightly different distribution, with a higher EpCAM signal relative to the others. A much greater deviation from the expected levels was observed in the H520 and A549 cell line with much lower specific binding.

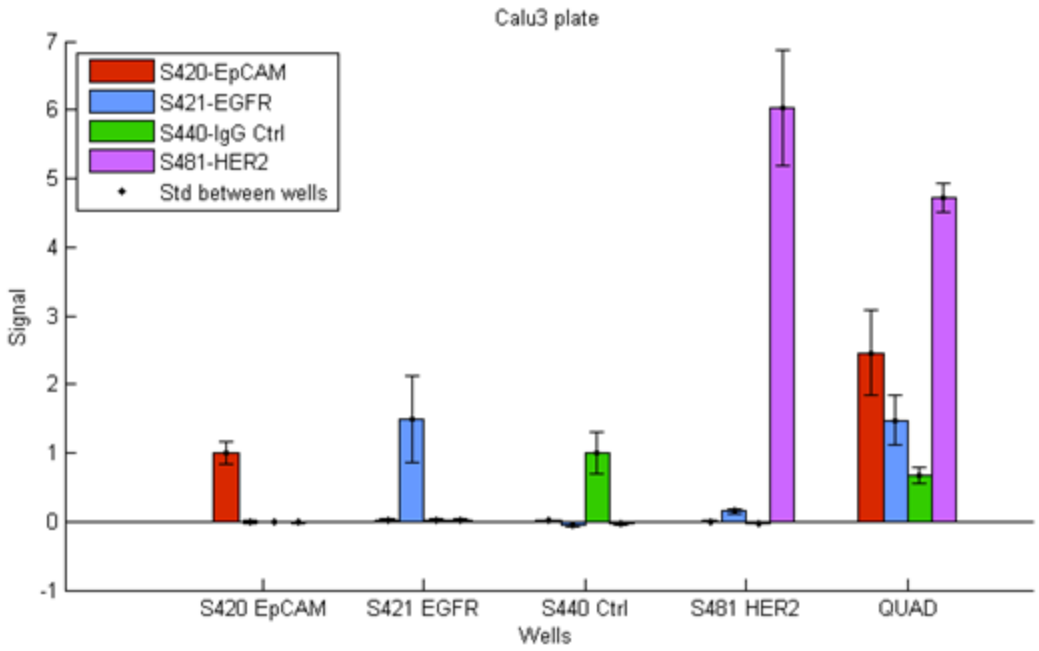

Figure 2.6 SERS NP binding distribution for the Calu-3 cell line. The first 4 wells were treated with a single type of SERS NP, as indicated on the x-axis. The final well, QUAD, was treated with an equimolar mixture of all 4 SERS NP types. (N=3 repeats, with bars indicating ±1 s.d.). The y-axis signal is the intensity value relative to the control sample.

2.8. *In Vitro* – Results In Suspension

When the cells are in an adhered state their antigen locations and availability might be changed from when they are not adherent in a single-cell suspension state, since the latter have more surface area available for binding. Since so little binding was observed in the adhered-cell state this was thought to help increase the amount of SERS NPs bound to the cells. Also the amount of cells in the sample was also increased to try to obtain more bound SERS NPs. Experiments in the suspension state were performed to have better binding and stronger signal.

The results presented here are with the MGH7 cell line. This is a lung cancer cell line also overexpressing EGFR like A549. The A549 presented very minimal EGF- conjugated SERS NP binding, so that this additional cell line was also tested. The A549 cell line results are not shown for in-suspension staining.

2.8.1. Flow Cytometry vs. SERS multiplexing (In Suspension)

To be able to determine if the SERS multiplexing is an accurate representation of the actual expression levels of receptors on the cell, flow cytometry was used as the gold standard for in-suspension experiments. The flow cytometry results represent the relative amount of binding of one type of antibody to its cell surface antigen compared to the other type of antibodies. This was done in a non-competitive manner, where each type was tested in separate tubes.

The expression levels of the antigens on the cells according to flow cytometry can be seen in **Figure 2.7** (bottom). The signal for each of the targeted antibodies is normalized to the signal from the untargeted binding. The Calu-3 cell line has a very high expression level of HER2 antigen and low expression of EpCAM and EGFR. The H520 cell line has a high expression of EpCAM and low expression of EGFR and HER2. The MGH7 cell line, similarly to A549, has high expression of EGFR and low expression of EpCAM and HER2.

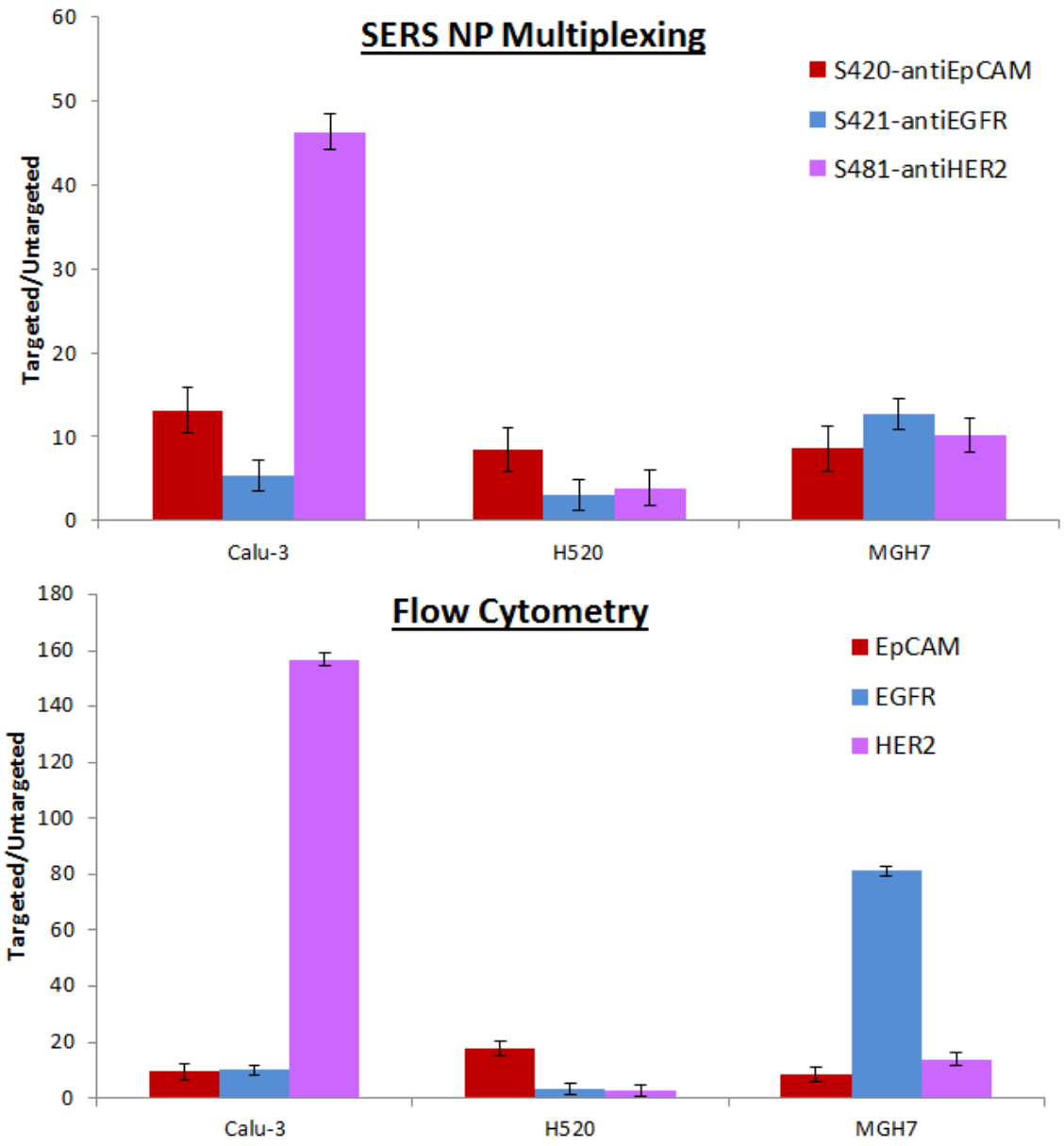

Figure 2.7 The binding of the SERS NPs compared to flow cytometry for three different cell lines. (N=3 repeats, with bars indicating ±1 standard deviation)

The Raman signal collected from the multiplexed spectrum, **Figure 2.7** (top), shows similar trends as the flow cytometry results, but is not identical. SERS multiplexing should be able to distinguish between low and high expression. The Calu-3 cell line exhibits the most similar trends, with the signal pertaining to the HER2 antigen being much stronger than for the other

types of antigens. The H520 cell line seems to trend towards stronger EpCAM signal relative to the other types of SERS NP binding. This is as expected from the flow cytometry results but is not definitive since the signal from the other SERS NP is not as strongly separated from it and the error bars are either overlapping or close to overlapping. The MGH7 cell line had the least definitive results, where the signal from the EGFR associated SERS NP had much less separation from other antibody binding than expected from flow cytometry.

2.8.2. Increasing Concentrations

If low staining concentrations are used, then the saturation of receptors on the cell may not be achieved. This could potentially cause the binding assay to be dependent on the properties of the binding SERS NPs and so not be an accurate representation of the expression levels on the cells. Hence, to determine whether the staining concentration is enough to reach saturation, increasing staining concentrations of 30-240pM were used to target two of the cell lines. The same staining procedure as above was used, varying the staining concentrations.

The Calu3 cell line (high HER2) and the H520 cell line (high EpCAM) showed increasing signal intensity with increasing staining concentration for each of the four targeted SERS NPs. This demonstrates that the SERS NP staining is not saturating even at maximum concentration of 240pM used in this experiment. While this is the case, the signal detected from the bound SERS NPs has the same trends as the expected results, where the SERS NPs signals being highest for the most highly expressed antigen binding. In **Figure 2.8**, the specific binding is normalized to the non-specific binding at each concentration and the points are fitted to a line to aid visualization.

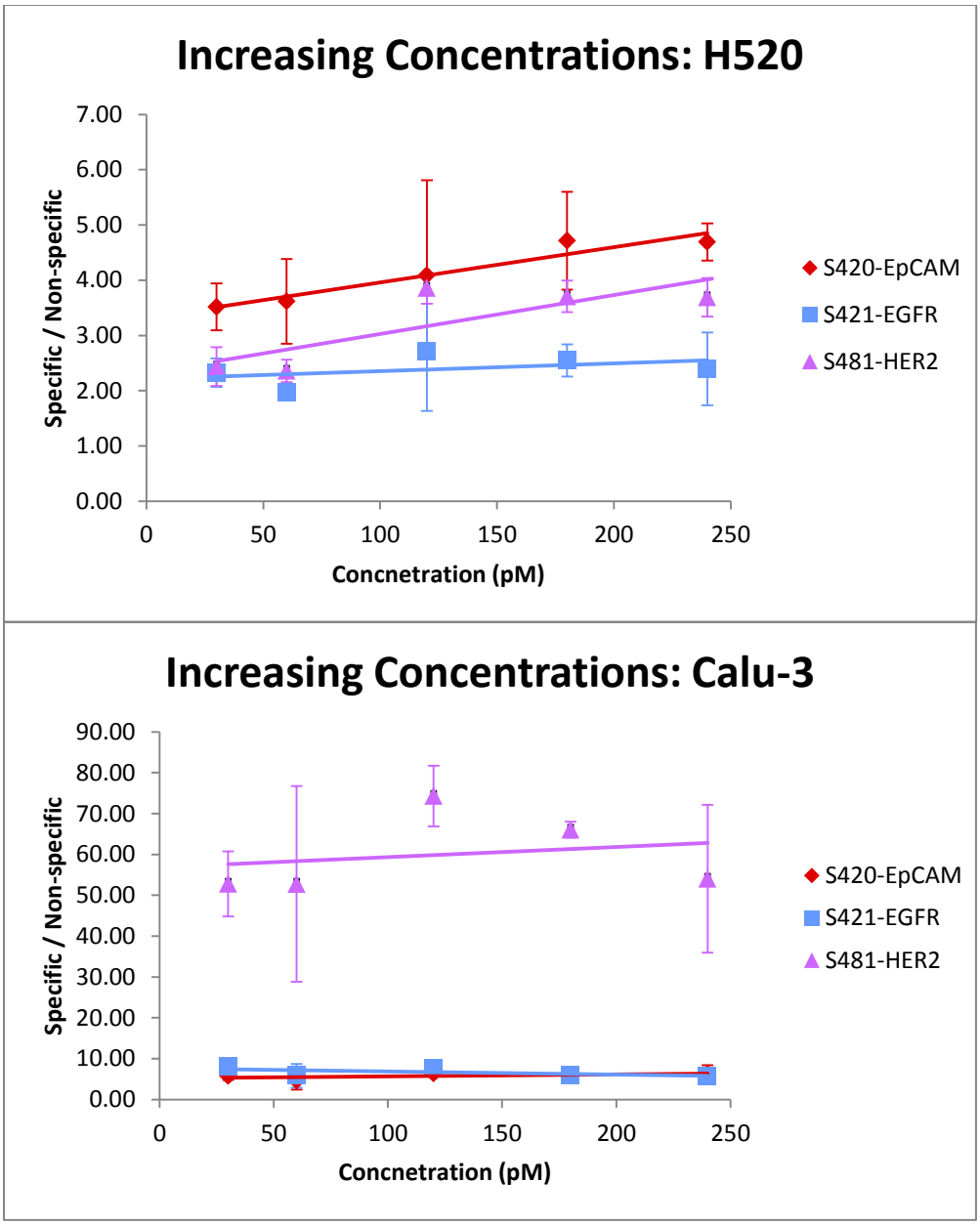

Figure 2.8 SERS NP binding with increasing staining concentration, normalized to the non-specific binding of the isotype control SERS NPs (N=3 acquisitions, 60s each, bar, ±1s.d.).

Exploring this further, after normalizing the calculated amount of targeted SERS NPs to the untargeted SERS NPs, the trends differed between the two cell lines. While the Calu-3 relative

amount stayed the same with increasing concentration, the H520 relative amount increased with higher concentrations for EGFR and EpCAM binding. This difference might be due to the H520 cells having a smaller total expression of the targeted antigens than Calu-3 or the antibodies binding differently in this cell line. According to the flow cytometry results, the total amount of signal from anti-HER IgG binding in the Calu-3 cell line is much higher than for any of the targeted binding for the H520 cells. The change in staining concentration will then affect the antibody binding in H520 more than in Calu-3 cells.

2.8.3. Blocking Experiment

To determine whether the SERS NPs are binding the cells specifically, an experiment was performed to try to first block the targeted sites, followed by staining with the SERS NPs. The cells were initially blocked with the same type of antibody as conjugated to the SERS NPs and then stained with a 40 pM concentration of each SERS NP in separate tubes and also in a quadruplex SERS NP staining mixture.

There was not much evidence of blocking found in the Calu-3 or H520 cell lines: **Figure 2.9**. There was some signal decrease in the Calu-3 cells in the anti-HER2 and anti-EGFR antibody binding. The H520 cell line showed the same signal from the targeted SERS NPs in both the blocked and unblocked cases.

The estimated amount of antibodies that were used for blocking is about 8×10^8 IgG per cell. This is about 5 orders-of-magnitude greater than the SERS NPs applied per cell, which was about 6×10^3. It is known that the SERS NP staining is not saturating the cells, so the blocking experiment would not perform as successfully, but some decrease in signal was still expected.

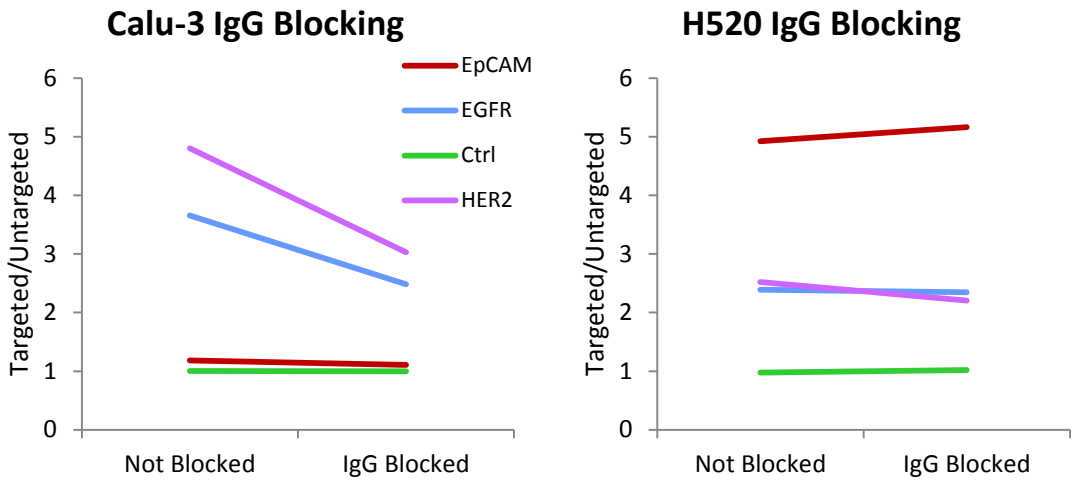

Figure 2.9 The binding of the targeted SERS NPs to Calu-3 and H520 cells before and after IgG blocking of the targeted sites (N=1).

2.8.4. NP per cell Calculation

An approximation was used to determine how many NPs were bound to the cells. The Raman signal collected from the cell solution was compared to the signal from the known concentration reference samples that were used for the least squares algorithm. The concentration of the cell suspension, C_{cells}, is directly proportional to the intensity of the Raman signal, I_{cells}, and to the scattering factor, s_{cells}, which depends on the type of cells used and the concentration of cells in suspension. The concentration is also inversely proportional to the acquisition time, t_{cells}, and the laser power, P_{cells}. Thus,

$$C_{cells} \propto \frac{I_{cells} \cdot s_{cells}}{t_{cells} \cdot P_{cells}} \tag{2.4}$$

To determine the concentration, the signal intensity was compared to that of the reference spectra of known concentration used in the least-squares calculations. Thus, the concentration of SERS NPs in the cell suspension was determined by

$$C_{cells} = C_{ref} \frac{t_{ref} P_{ref}}{I_{ref}} \cdot \frac{I_{cells} s_{cells}}{t_{cells} P_{cells}} \quad (2.5)$$

, where the ‘ref’ subscript refers to the corresponding values for the reference spectra.

To confirm the state of the Raman spectrometer and account for any changes in signal intensity, the calibration spectrum of a silicon sample was collected before each day’s measurements. Changes could occur from fluctuations in laser power or the instrument collection efficiency. The Raman peak from the silicon sample was used as an indirect measurement of the laser power and signal collection.

Light scattering by the cells reduces the backscattered Raman that is collected by the system. This must be accounted for each of the cells used, as it will depend on their size and shape. To estimate the scattering coefficient for the system, known low concentrations of SERS NPs (1pM) were combined in suspension with known cell concentrations, ranging from 0 to 1.5 million in 1mL. This was repeated in triplicate. The reduction in the collected signal can be seen in **Figure 2.10**, where the signal at 1.5 M/ml cell concentration decreased 5-fold. The curve of the scattering coefficient was modeled by a double-exponential decay from the intensity of the null suspension. This represents the decay experienced by both the incident light and the backscattered light due to scattering in the sample.

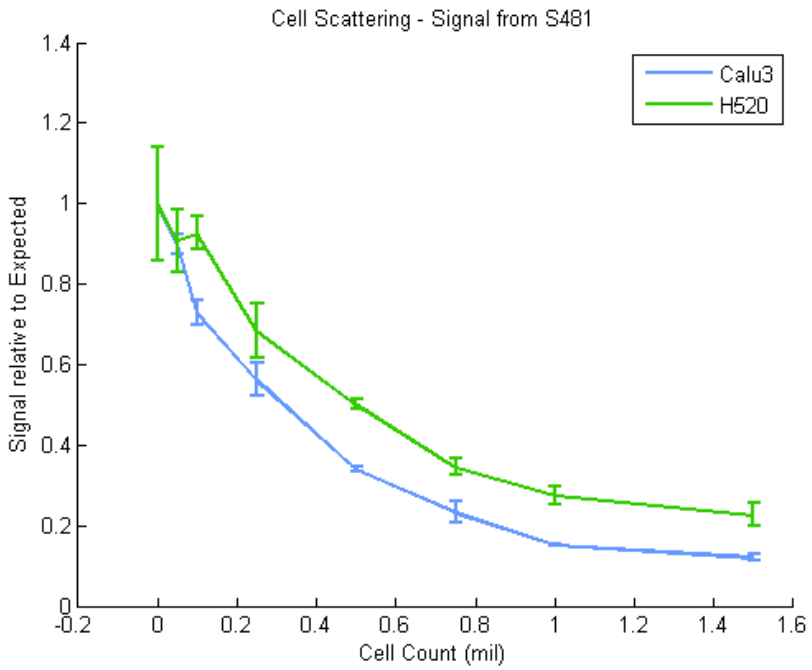

Figure 2.10 The decrease in SERS signal due to light scattering by cells, illustrated here for the S481 reporter, as the number of cells increases in the suspension.

This effect is intensified due to the long focal length of the chosen lens for the system. The rate of signal decay would be less if a shorter focal length lens was used, since the distance traveled by the light, and thus interactions with scatterers, would be reduced.

This concentration approximation is then used to calculate the amount of SERS NPs per cell:

$$N_{per\ cell} = \frac{C_{cells} \cdot V \cdot N_{Avogadro}}{N_{cells}} \tag{2.6}$$

, where $N_{per\ cell}$, is the number of NPs per cell; V is the volume of the suspension; $N_{Avogadro}$ is Avogadro's number and N_{cells} is the number of cells in suspension.

The number of bound SERS NPs per cell can be estimated using this approximation: in general, this number is around 10^3, which is lower than the copy number of antigens expected, where overexpression can be around 10^5 HER2 antigens per cell [96].

2.9. Discussion

The conjugation of the antibodies to the SERS NPs was successful in the sense that the conjugates were stable and reproducible. The conjugation gave a product that was consistent between the different antibodies used, and yielding approximately the same concentration of antibody per SERS NP, which allowed for direct comparison of the SERS signal and the binding of the NP.

The difficulty came in selecting an antibody that is compatible with the conjugation method. In the case of these experiments, the purchased antibody needed to be chosen based on a set of criteria that limited the selection. First, the antibodies needed to be monoclonal in order to achieve a consistent distribution of the same type of antibody surrounding the SERS NP and so increase the reproducibility. Second, the antibody had to be from the same species as the other chosen antibodies so that comparisons of the binding could be made using secondary fluorescently-labeled antibodies. Changing the host species of the antibody will affect the type of secondary antibody used and the binding properties to the primary antibody. Third, the antibody needed to be stable and be stored in a non-protein based solution. This is the most important condition; if there are any other types of protein in the antibody stock solution then that protein will be conjugated to the SERS NPs via an acylation reaction, just as the antibody would. A common protein used for stability is BSA and this is found in a much higher concentration than the antibody in the stock solution, so that the SERS NP would be covered in a much higher proportion of BSA than the antibody, thereby making the product unusable for targeting. Potentially, the protein used for stability could be extracted using purification spin columns. This has been testing in our group previously and we found that this was not a successful method. Considering all these selection criteria, the amount of antibodies that can be used for synthesis to the SERS NPs is limited. Thus, for these experiments only antibodies matching these conditions were considered.

Evaluating the results obtained from both the in-suspension and plate-staining of the cells, it was found that the SERS NPs were not binding as expected. The most likely explanation for lack of binding for some of the conjugated antibodies is the choice of antibody itself. The conjugation reaction occurs with the lysine residue on the antibody. This residue is positively charged at physiological pH, so that is found most commonly on the outside surface of the protein and the site is available for conjugation. Considering that the antibody binds to antigens mainly via electrostatic forces from its residues on the outside surface, it is plausible that the lysine residue used to conjugate the antibody to the SERS NP is also necessary for binding to its target or other nearby residues. Decrease in binding affinity of antibodies after conjugation via amine groups has been reported in several instances in the literature. For example, Trubetskoy *et al.* [97] found that, when conjugating an antibody using a lysine group to poly[l-lysine] (PLL), the affinity decreased due to a blockage of the antibody-binding site. They overcame this problem by using a carbohydrate-directed conjugation to target the conjugation towards the hydroxyl group of the antibody. Vira *et al.* [98] describe the avidity or the combined effect of all the binding interactions (affinities) of antibodies before and after conjugation to a fluorophore. Fluorescein isothiocyanate (FITC) was shown to decrease the avidity of Fc125 antibody, against hemagglutinin. This was also observed in Trastuzumab conjugation to AlexaFluor 647 [99]. However, fluorophore conjugation is different from the SERS NP conjugation. Fluorophores are small relative to the antibody so that the product of the conjugation procedure will have many fluorophores on the antibody surface, whereas SERS NPs are many times larger than the antibodies so that there are many antibodies conjugated to the NP surface. This difference does not impact the conclusion that the binding site might be affected due to the lysine chemistry.

Although there might be even more accurate ways to determine the antigen numbers on the cell surface, flow cytometry does seem applicable here, since it represents a similar way of binding of the antibodies as the SERS NP staining. The antibodies were used at the same relative concentration to be able to compare it to the SERS NP staining, which were also taken at the same relative concentration to each other. The technique was used to assess how the SERS NPs affected the binding of the antibodies.

The only difference between the staining methodology between the two techniques is that the flow cytometry staining was done non-competitively (in separate tubes), while the SERS NPs

were stained under competitive conditions (all for in the same tube). This distinction is not as important under non-saturation conditions.

While the staining concentration used in this work does not saturate the receptors, the ratios of the targeted SERS NP binding show the same trends as the expected results. It would be interesting to perform a full saturation experiment to determine how many NPs are needed. However, this is not practical as it would require the use of a large amount of Au NPs and this is not a cost-wise realistic. Thus, the SERS NPs were testing only in under-saturated conditions and equimolar staining concentrations as their expected use.

A better way of comparing the classification abilities of this technique would be to perform a titration experiment which saturates the receptors since there are differences between antibodies. This would account for the binding affinity differences between the antibodies and determine if the binding is accurately reflecting the expression levels of the antigens.

2.10. Summary

Claiming quantifiable results using SERS multiplexing requires thorough investigation of the NPs binding. The NPs need to be comparable to each other to support a claim that the signal collected means that the corresponding antigen is overexpressed in the cells. It was concluded from the above experiments that the SERS NPs were successfully conjugated, by showing that there was no aggregation by evaluating the Au absorption peak. The relative amount of antibodies conjugated to the SERS NPs when comparing the four different types was similar. The least-squares algorithm was shown to be an effective method for spectrally unmixing the Raman signals to determine the relative amount of SERS NPs present in the mixture.

Overall, the SERS NP staining shows different results than the flow cytometry gold standard. The most highly expressed antigen can be identified in the Calu-3 cell line with HER2 and less convincingly in the H520 cell line with EpCAM. The binding of the antibodies was most likely affected by the conjugation process to the SERS NP. This can be seen by the change in total distribution of the targeted SERS NP binding compared to the flow cytometry results. There is

not as much separation between the relative amounts of the most highly expressed antigen and the others.

Chapter 3
Intrinsic Raman Spectroscopy

3.1. Introduction

Optical imaging has the advantage of having very high resolution with potentially high sensitivity and specificity. While tissue imaging modalities such as optical coherence tomography (OCT) and autofluorescence imaging provide morphological or functional information about the tissue, they fail to provide any biochemical information that is important in identifying any early neoplastic changes.

Raman spectroscopy of cells or tissue is a powerful technique when combined with multivariate statistical analysis for cell and tissue characterization. The small changes in composition due to changes in chemical bonds can be detected in the Raman spectra. They provide a sort of 'molecular fingerprint', enabling classification of the cells or tissue as normal or diseased. In optical microscopy this technique has been used in many single-cell analyses, for example, classifying human breast epithelial cells [42] and leukemia cells [44] and for differentiating between stem cells and their differentiated progeny [100].

This work focuses on ovarian cancer and validating the use of Raman spectroscopy to classify cells as either normal or malignant. It serves as the first step in being able to translate this technology to identifying ovarian cancer cells in tissue, either for (1) screening, (2) surgical margin determination or (3) treatment response monitoring.

To model ovarian cancer *in vitro*, two cell lines were used - mouse ovarian surface epithelium (MOSE: M0505) cells and spontaneously-transformed surface epithelium (STOSE) cells. STOSE cells have been identified as the first spontaneous murine model of HGSC (high-grade serous ovarian cancer) [101], which is the most common subtype for ovarian cancer. Raman microspectroscopy along with multivariate analysis was applied to these two types of cells for the first time.

The origins of ovarian cancer are not yet fully resolved. There is controversy as to whether the cells are of ovarian origin or fallopian tube origin [100,102,103]. This is difficult to determine in patients, since most often the disease is diagnosed when the tumor has progressed to later stages and occupies the whole abdominal cavity region. The original paper describing the STOSE cells argues that there is validity in thinking that some ovarian cancer originates in the ovarian surface epithelial cells, since the cells were observed to originate from normal ovarian surface epithelium as shown in mice [101].

3.2. Acknowledgements

The work in this chapter focuses on the application of Raman spectroscopy to analyze spectra collected from live cells. This is an initial step in developing a method to detect the differences that occur when cells transform from normal to malignant. This chapter is a summary of a published paper: S. Borel, E. A. Prikryl, N. H. Vuong, J. Jonkman, B. Vanderhyden, B. C. Wilson and S. Murugkar, Discrimination of normal and malignant mouse ovarian surface epithelial cells in vitro using Raman microspectroscopy, *Analytical Methods* 7, 9520-9528, 2015 [104]. My contribution to this work involved designing the live-cell experimental protocol, performing the Raman signal collection and contributing to the development of the spectral analysis methodology. The multivariate statistical analysis was performed by Emil Prikryl and Dr. Sangeeta Murugkar at Carleton University and the cell lines were provided by Dr. Barbara Vanderhyden's group at the University of Ottawa.

3.3. Methods

3.3.1. Cell Culture

MOSE and STOSE cells were cultured in MOSE media according to the protocol described in Gamwell et al. [105]. After the cells reached confluence, TrypLe (Life Technologies, Burlington, ON) was used to detach them from the tissue culture dish. They were then plated at low density on a thin 25.4 mm diameter quartz coverslip (01019T-AB SPI Supplies, West Chester, PA) contained in a 60 mm tissue culture dish.

It had been demonstrated in earlier work [51,106] that both cell cycle progression and changes in cell culture confluency, are major sources of spectral variability in Raman spectra of single cells measured from the same cell culture. Since the objective of this work was to test if Raman spectroscopy can detect biochemical differences between MOSE and STOSE cells, appropriate steps were implemented, as follows, to minimize the effect of confluency and the cell cycle on the measured spectra. Both MOSE and STOSE cells were propagated until they were 95% confluent and a cell-cycle inhibitor, Roscovitine (CST, Whitby, ON, Canada), was used to arrest the cells in G1/G0 and G2/M phases. A solution of Roscovitine was added to the cell media to a final concentration of 20 μM in 0.1% DMSO. This concentration was determined from a set of preliminary dose-response experiments to identify the maximally-effective concentration that inhibits cell proliferation, and was used previously [107] without apparent cytotoxicity. The same amount (0.1%) of DMSO was added to the untreated control cells for which the Roscovitine was not used. This amount of DMSO is expected to have no impact on cell viability.

Imaging was performed 24h after Roscovitine treatment for cell cycle arrest. The quartz coverslip containing the adhered cells was transferred to a Chamlide magnetic chamber (Live Cell Instrument, Seoul, Korea) and filled with fresh media containing either Roscovitine in DMSO or DMSO only. A Chamlide stage top incubator (Live Cell Instrument, Seoul, Korea) was used to maintain the cells under optimum conditions of 5% CO_2, 50% humidity and 37 ^{o}C during the acquisition of Raman spectra.

3.3.2. Raman Spectra Collection

Raman spectroscopy measurements were made using a confocal Raman system (Renishaw inVia, Hoffman Estates, IL, USA) as shown in **Figure 3.1**, configured to an inverted microscope base (Leica DMI6000B, Concord, ON, Canada) fitted with a 63X water-immersion objective lens. Excitation light of 785 nm was used at a power of 70 mW at the sample. Each spectrum was acquired with a total integration time of 1 min (6 scans each of 10s).

Figure 3.1 The Renishaw inVia Raman system equipped with a Leica microscope (left) and a Renishaw spectrometer (right).

Spectra were acquired from 40 randomly, manually selected cells adhered to the quartz substrate for each of the four groups (MOSE control, MOSE treated, STOSE control, and STOSE treated) for a total of 160 spectra. **Figure 3.2f** shows a typical image of MOSE cells in the field of view of the Raman microscope obtained using Differential Interference Contrast (DIC). The image illustrates the laser beam superimposed on one MOSE cell such that the sampling area is centered on the nucleolus and spans mainly the nucleus with minimal contribution from the cytoplasm. The laser beam has a 2 µm by 15 µm focus spot on the sample. The laser beam originates as an elongated shape. A circular laser spot can be obtained by using a pinhole but this results in a 95% decrease in laser power. Since the Raman signal is intrinsically weak, special resolution was sacrificed for a higher signal-to-noise ratio.

Background spectra were acquired and averaged over four different locations of the quartz substrate containing only the cell media with or without Roscovitine in DMSO.

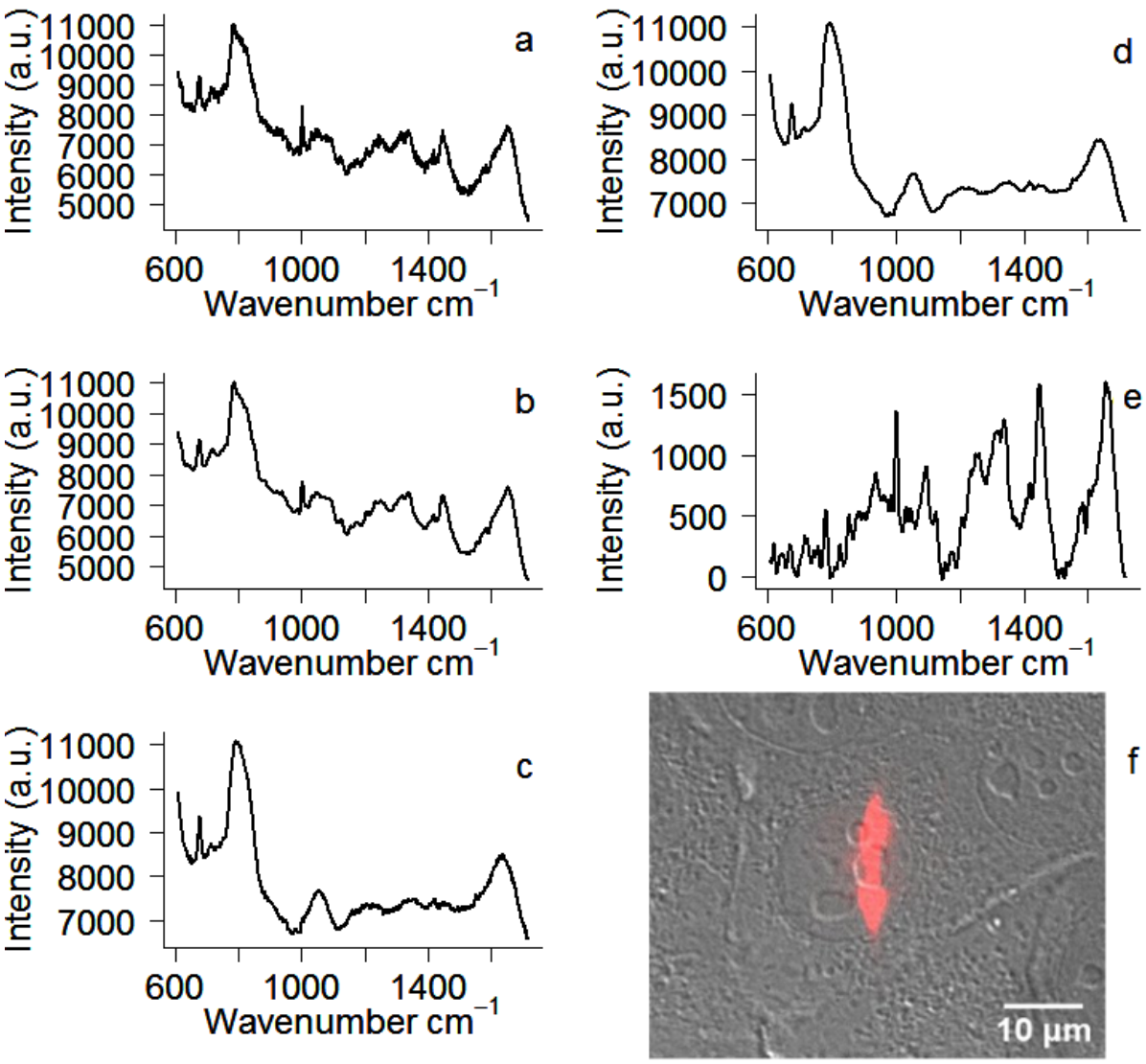

Figure 3.2 Pre-processing steps for a sample Raman spectrum in the cell nucleus, where the x-axis represents the wavenumber and the y-axis represents the signal intensity. (a) Raw cell spectrum, (b) with Savitzky–Golay smoothing, (c) average of 4 raw background spectra, (d) Savitzky–Golay smoothing applied to background spectrum, (e) correction for the background contributions using the SMIRF background removal algorithm yielding the extracted cell spectrum, and (f) the 785 nm laser focus spot (2µm x 15µm) superimposed on a differential interference contrast (DIC) image of the cell monolayer, acquired with a 63x water immersion

objective. The laser beam is centered on a nucleolus of the cell and spans the nucleus. (Reprinted with permission from ref. [104], Copyright 2015 The Royal Society of Chemistry)

3.3.3. Data Processing and Multivariate Analysis

3.3.3.1. Spectra processing

Raw Raman spectra of the cells contain many confounding factors and need to be processed before they are used for analysis. The cell spectrum is confounded by noise and background components, including from the quartz, media, media additives and fluorescence signal. The spectra also need to be smoothed to reduce noise, followed by a smoothed background spectrum subtraction to obtain the Raman spectral components originating from the cell. This section will give a detailed explanation of the steps in this process to obtain the pure cell spectrum and reduce the variability in the collected Raman signal of the cells for meaningful comparisons.

First, a Savitzky-Golay smoothing filter (using 15 spectral points and second-order polynomial) was applied to each raw spectrum and the collected background spectrum. A background subtraction method was used to yield the true Raman spectra of the cells and reduce or eliminate the varying levels of background fluorescence signal from the cells, media, DMSO and Roscovitine and the Raman signal from the quartz substrate. These background contributions were removed using the Spectrum-based Method for Iterative Removal of Fluorescence (SMIRF) background removal algorithm developed in MATLAB® by Beier and Berger [108]. The average of the smoothed background spectra collected from the cell media was used as the 'contaminant', or components apart from the cell, and a 5th-order polynomial was used to model the broad and slowly-varying background fluorescence. SMIRF uses an iterative algorithm to best fit these components to the raw spectrum and obtain the Raman components of the cell.

The remaining variability between spectra is the signal intensity that may fluctuate due to the amount of biological sample being sampled by the laser beam. This may be caused by the slight differences in cell morphology (shape/orientation) and in focus height. Each spectrum was then normalized to the total area under the curve to produce the final spectrum: **Figure 3.2e**. Other methods have used the area under the peak at 1450 cm^{-1}, originating from CH deformation, to normalize the spectra, but this then makes the assumption that it is representative of the amount

of biological material sampled [109,110]. Normalizing to the full area under the curve method makes no assumptions about the cell biochemistry. Due to minute wavenumber shifts in the system, the Raman wavenumbers were then mean-centered for subsequent processing and comparison.

The quartz Raman contribution to the background spectra was significant and had fluctuating signal intensity depending on the location, likely originating from inconsistent thickness of the quartz coverslip. Hence, the background spectra were stronger than the cell spectra in certain spectral ranges, most notably in the 800 cm^{-1} region. In addition, a sharp peak at 670 cm^{-1} from DMSO [111] may not have been removed consistently and could result in artificial contributions in this region. DMSO also produces a small and broader peak at 710 cm^{-1} but this peak has less impact on the background subtraction because of its lower intensity.

3.3.3.2. Multivariate Statistical Analysis

Principal component analysis (PCA) and linear discriminant analysis (LDA) were chosen for classifying the groups of spectra. These are two complementary multivariate statistical techniques that separate the spectra based on the statistically important factors in the data. They reduce extraneous and redundant factors, which improves the accuracy of the classification.

PCA is an unsupervised technique that is used commonly to reduce the dimensionality of spectroscopic data. It computes new variables that represent the data set, called principal components (PCs). These maximize the variance of the data set and are orthogonal to each other [112]. The first PC accounts for the most variance in the data set, with diminishing contributions from the following PCs. The most important spectral features that represent the most difference between the spectra in the data set are emphasized and the redundant information is reduced.

Linear discriminant analysis (LDA) uses the PCs to maximize the separation between the two groups in the data set and minimize the variation within the groups [113]. It computes the component, called the linear discriminant, along which the separation between the two groups is greatest. The LD loading plot can be examined to view the separation that is determined to exist between the two groups using LDA.

In this work, the two groups to be separated using the PCA/LDA method were the spectra from the MOSE and the spectra from the STOSE cells. PCA combined with LDA has been widely used to classify and analyze Raman spectra in cells and tissues [39,43,114-118].

PCA was performed on the cell spectra of the control and treatment MOSE and STOSE groups using the STATS package in the R programming language [119]. LDA was applied to the output of PCA algorithm using the MASS package in R [120]. The choice of the number of PCs to use in the LDA was determined by examining the location of the elbow of the scree plot and keeping the principal components that precede the elbow. From the scree plot, only the first two PCs were found to be the dominant contributors to the LDA, as explained below.

To evaluate the effectiveness of this method for the separation of the two groups, the specificity and sensitivity of the classification for the confusion matrix were computed taking the STOSE classification as the positive and the MOSE classification as the negative. This was based on the well-established leave-one-out-classification validation (LOOCV) method [121]. The loadings of the first principal component and the linear discriminant were examined to identify those Raman peaks that accounted for the biomolecular differences between the cell types.

3.4. Results

3.4.1. Untreated Control Samples

The average Raman spectra from 40 MOSE and STOSE cells collected can be seen in Fig **Figure 3.3**. The Raman peaks for these spectra represent the superposition of peaks mainly from the cell nucleus containing protein, lipid and nucleic acid, and contributions that may be from polysaccharides. The likely molecular assignments for the peaks can be found in **Table 3.1**. The peaks are considered as a wavenumber range centered at the values listed.

The wavenumbers, or the x-axis, represent the Raman shift of the scattered light from the cells. The peak locations for the average spectra of the two cell types are the same, but the intensities differ: **Figure 3.3**. This is because the composition of the cells does not change markedly between the cell lines, but the relative amounts molecules do differ, as represented by the

intensity changes between spectra. In the difference spectrum, **Figure 3.3c**, the positive bands represent an increased intensity in the STOSE cells compared to the MOSE, indicating an increased concentration of that constituent. The peaks marked 1 and 4 in the difference spectrum indicate higher nucleic acid content in the STOSE cells compared to the MOSE. The specific peak at 666-678 cm^{-1} (1) is associated with the guanine ring breathing modes in DNA/RNA bases and C-S stretching mode in cytosine [122,123]. The band at 766-786 cm^{-1} (4) is associated with uracil, thymine and cytosine ring breathing modes in DNA/RNA bases. For certain wavenumbers there is an overlap of peaks that originate for different types of components. The mixed assignment peaks, marked in grey, can also contribute to increased nucleic acid content in the STOSE cells. The peak at 748–753 cm^{-1} can be from the ring breathing mode in thymine in DNA or symmetric breathing of the amino acid tryptophan associated with proteins. The peak at 1415–1421 cm^{-1} (6) is associated with adenine, guanine and deoxyribose in DNA but could also have contributions from C-H deformations in proteins or lipids.

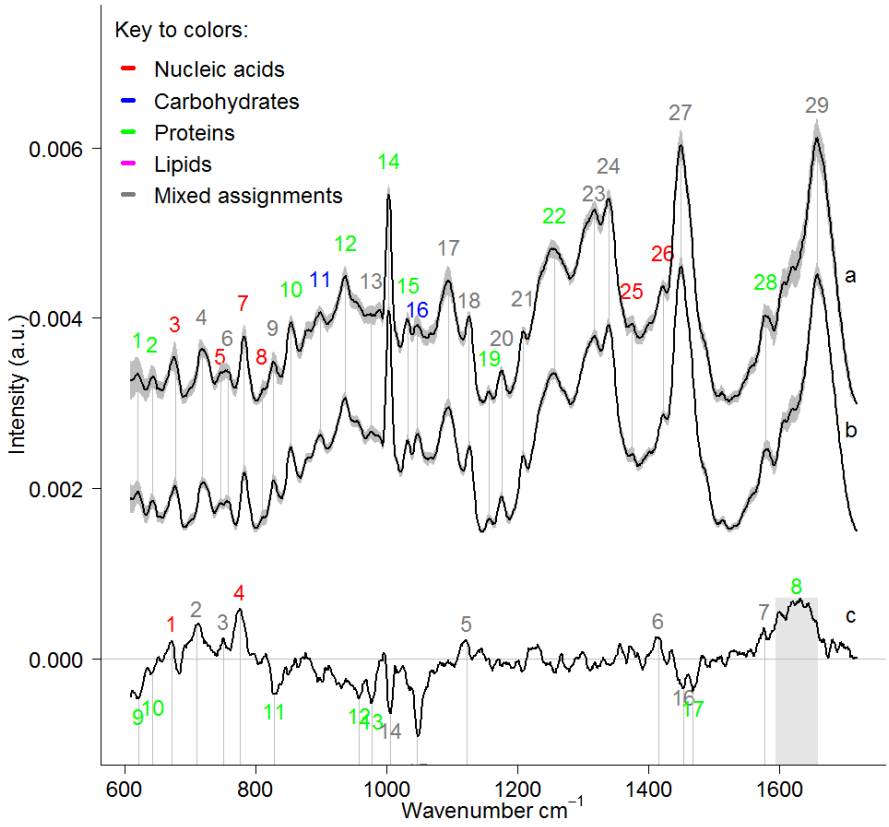

Figure 3.3 (a) Average STOSE spectrum, (b) average MOSE spectrum and (c) difference spectrum generated by subtracting the average MOSE spectrum from the average STOSE spectrum in non-arrested cells. The grey curves are the standard-deviation envelopes. Spectral peaks of MOSE and STOSE spectra are highlighted by grey lines and annotated with numbers assigned as per **Table 3.1**. The spectra are shifted in the y-axis for visual clarity. (Reprinted with permission from ref. [104], Copyright 2015 The Royal Society of Chemistry)

Band	Wavenumber (cm^{-1})	Molecular Assignment(s) with Wavenumbers from Literature
1	618 621	C-C twisting (p) C-C twisting mode of phenylalanine (p)
2	640 643	C-S stretching & C-C twisting of proteins - tyrosine (p) C-C twisting mode of tyrosine (p)
3	678	**G ring breathing (d)**
4	717-719	Choline (l), C-N (membrane phospholipids head) (l), **A (d)**
5	746	**T (ring breathing mode of DNA/RNA bases) (d)**
6	755 759	Symmetric breathing of tryptophan (p) Tryptophan (p), ethanolamine group (l), or phosphatidylethanolamine (l)
7	782	**U,C,T ring breathing (d)**
8	810	**phosphodiester (d)**
9	827	Out-of-plane ring breathing, tyrosine (p), **O-P-O stretch (d)**
10	853	Ring breathing mode of tyrosine & C-C stretch of proline ring (p)
11	898	Monosaccharides (beta-glucose) and disaccharides, (C-O-C) skeletal mode (c)
12	936 937	C-C symmetric stretch backbone, α-helix (p) C-C backbone (p)
13	974 980 980	**Ribose vibration (d)** C-C stretching beta-sheet (p) =CH bending (l)
14	1000/1/3	Phenylalanine (p)
15	1030 to 1040	C-H bending phenylalanine (p)
16	1048 1049	glycogen (c) C-O stretch in carbohydrates (c)
17	1093 1094 1094	**Symmetric O-P-O stretch of DNA backbone (d)**, or C-N stretch (p) **(d)** C-N stretch (p), or chain C-C stretch (l)
18	1125	C-N stretching (p); C-C stretching (l)
19	1158	C-C/C-N stretch (p)
20	1175/6 1176	**C, G (d)** C-H bending tyrosine (p)
21	1208	tryptophan, phenylalanyne (p), or **A, T ring-breathing (d)**
22	1254	C-N in-plane stretch (p)
23	1320 1320	**G (d)** C-H deformation (p)
24	1336-1345	**A, G (d**), C-H deformations (p), CH2 twist (l)
25	1373 1374	**T, A, G (ring breathing of DNA/RNA) (d)** **T (d)**
26	1421 1422	**A, G (d)** **Deoxyribose (d)**
27	1450	CH deformations (p, l)
28	1582	phenylalanine (p)
29	1656 1655-1680	amide (p) or C=C stretch (l) Amide I (p)

[a] Abbreviations: (d) nucleic acid, (c) carbohydrate, (p) protein, (l) lipid (G) guanine, (C) cytosine, (A) adenine, (T) thymine, (U) uracil

Table 3.1 Provisional molecular assignments for peaks in average MOSE and STOSE spectra. Bold indicates nucleic acid-related peaks. References: [109,122,52,124,125] (Adapted with permission from ref. [104], Copyright 2015 The Royal Society of Chemistry)

A paper on the STOSE cells by McCloskey *et al.* [101] guided the interpretation of the difference spectra and the PCA/LDA results. This paper describes how the STOSE cells differ from the original ovarian epithelial cells (MOSE) and how they are a model for murine HGSC. The STOSE cells have a doubling time of 13h, which is nearly 4 times faster than the MOSE cells. STOSE cells also have a higher degree of aneuploidy, or abnormal number of chromosomes, that is higher than observed in the MOSE cells. STOSE cells are mostly near-triploid and some polyploid, while MOSE cells are partly near-tetraploid and mostly near-diploid. From cell cycle analysis, a much higher proportion of STOSE cells were found to be in the S phase (45%) compared to the MOSE cells (10%). The STOSE cells have a total proportion of cells with higher DNA content in phases S and G2 that is over 50% higher than in the MOSE cells.

As expected, we observed higher peaks related to nucleic acid content in the STOSE cells, although this is not fully conclusive as there was some peak overlap with other components in the mixed peaks.

The PCA/LDA analysis of the cell spectra revealed separation of the two untreated groups as can be seen in the PC plot in **Figure 3.4a**. Two outliers were removed from the data set as evaluation of the spectra revealed that the background spectrum was not adequately removed. PC1 in large part represents the intergroup variability, while PC2 represents the intragroup variability. The first two PCs were used for the LDA analysis, as determined by the elbow of the scree plot that compares the contribution to variance of each of the components. This is where the component becomes relatively small and similar to the remaining components. Increasing the amount of PCs included in the LDA analysis will increase the classification accuracy but will be sample-dependent. The higher number PCs are due to subtle cell variations and noise contributions. They will not be informative for larger sample sizes. The accuracy that was achieved with the control samples was 72% (**Table 3.2**), with a sensitivity and specificity both of 72%. This limited accuracy of the technique was due to large intra-group variations that can in part be due to the cells being in different phases of the cell cycle. The results of the arrested cell cycle are described in the following sections.

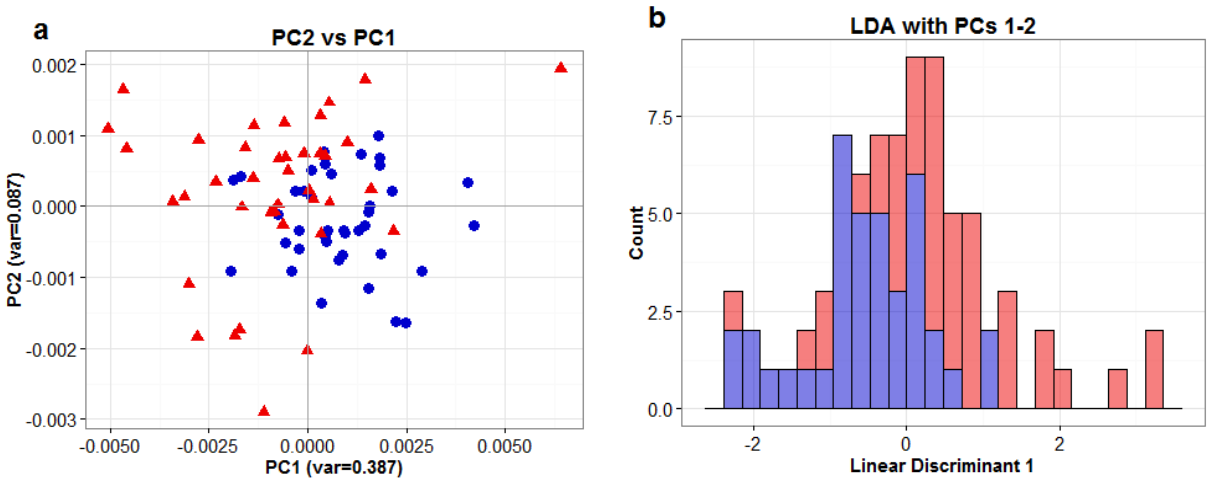

Figure 3.4 PCA score plots (a) and LDA histograms (b) for MOSE (blue) and STOSE (red) non-treated cells. (Reprinted with permission from ref. [104], Copyright 2015 The Royal Society of Chemistry)

Treated	Predicted		
Actual		MOSE	STOSE
	MOSE	34	6
	STOSE	3	37
	Sensitivity	92%	
	Specificity	85%	
	Accuracy	89%	
Control	Predicted		
Actual		MOSE	STOSE
	MOSE	28	11
	STOSE	11	28
	Sensitivity	72%	
	Specificity	72%	
	Accuracy	72%	

Table 3.2 Confusion matrix for LOOCV (leave-one-out-classification validation) on PCA–LDA of treated and control cells. (Reprinted with permission from ref. [104], Copyright 2015 The Royal Society of Chemistry)

3.4.2. Treated Samples

In order to extract biochemical differences between the two cell lines, differences in the cell cycle need to be controlled, so that the cells were treated with Roscovitine to arrest them in the G1/G0 and G2/M phases [126,127]. This was expected to reduce intra-group variability due to cell cycle differences in the proportion of cell found in the S phase.

The average of the collected spectra of 40 treated cells for MOSE and STOSE can be seen in **Figure 3.5a**. The spectra again look visually similar with no large differences, requiring quantitative analysis. The PCA analysis showed that the first two PCs accounted for a large portion of the variation seen in the data. There was much better clustering of the two groups, as in **Figure 3.5c**, compared to the first two PCs in the untreated group. PC1 accounted for more of the inter-group variation and 52% of total variation in the data. PC2 accounted for more of the intra-group variation and 6% of total variation in the data. The LDA analysis using the first two PCs showed better separation of the MOSE and STOSE groups. The sensitivity, specificity and accuracy were improved from the control group to 92%, 85% and 89%, respectively (**Table 3.2**).

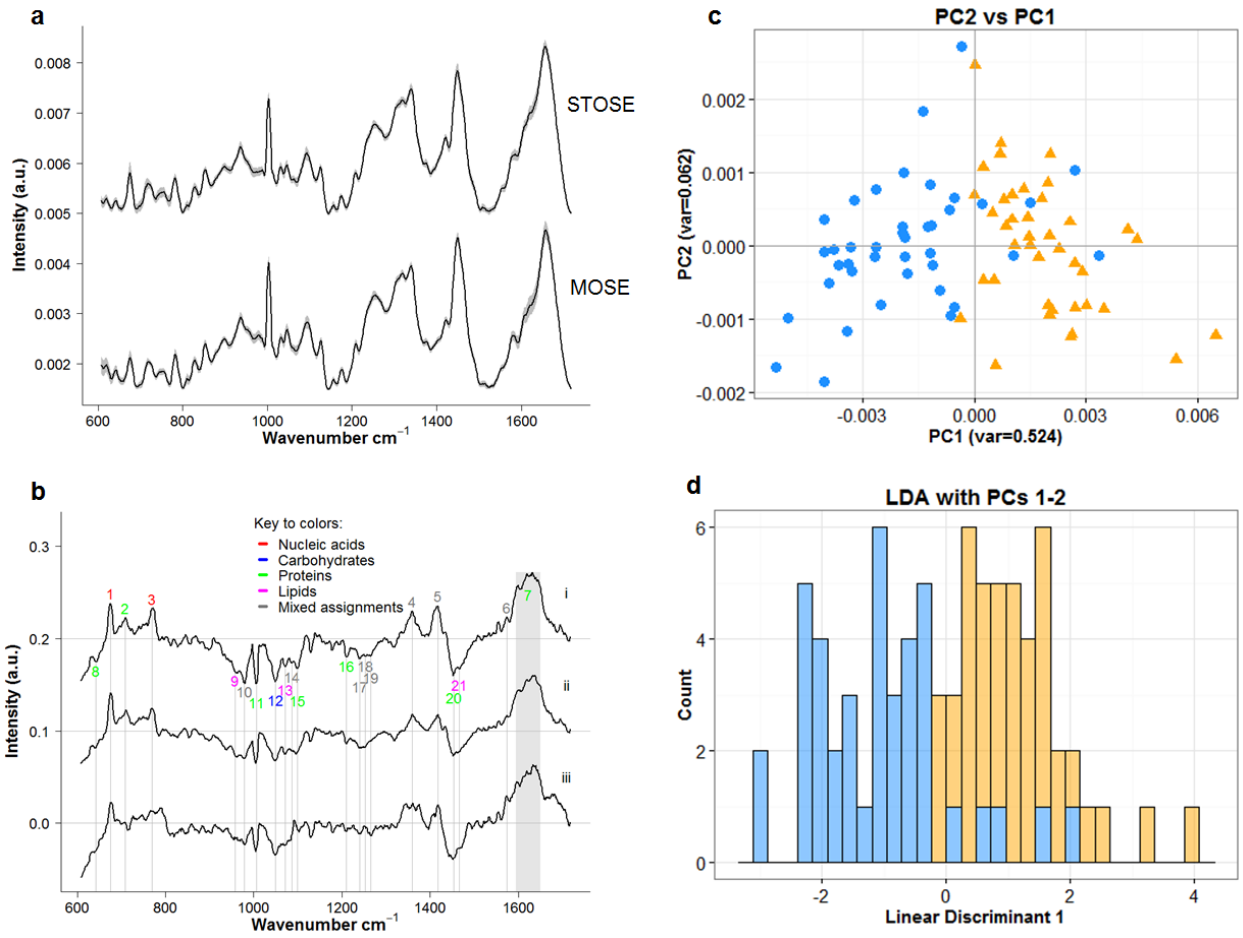

Figure 3.5 (a) Average STOSE and MOSE spectra from treated group with the gray shading indicating standard deviation envelopes. (b) Comparison of (i) difference spectrum (STOSE–MOSE), (ii) PC1 loadings plot and (iii) LD loadings plot for treated spectra. Spectral peaks are numbered and highlighted by grey lines, as per **Table 3.3**. (c) PC scores plots and (d) LDA histogram for treated MOSE (blue) and STOSE (orange) cells. (Reprinted with permission from ref. [104], Copyright 2015 The Royal Society of Chemistry)

Band	Wavenumber (cm ^{1})	Molecular assignment(s)	Band	Wavenumber (cm ^{1})	Molecular assignment(s)
1	653–673	**G (d)**	12	1048	Glycogen (c)
	678	**G ring breathing (d)**			
2	700–45	C–S trans vibration (aminoacid methionine) (p)	13	1073	Triglycerides (l)
3	766	**Pyrimidine ring breathing (d)**	14	1087–1090	C–C stretch, phosphate stretch (d,l)
4	1355/7	**G (d)**	15	1099	C–N stretch (p)
	1359	**Tryptophan (p)**			
	1360	**Tryptophan (p)**			
	1361/2/3/5	G (d)			
5	1417	C] C stretch in quinoid ring	16	1209	Tryptophan, phenylalanine (p)
	1420	CH_2 (p, l)		1210	Tyrosine, phenylalanine (p)
	1420/1	**Deoxyribose (d)**			
6	1573	**A, G (d) , tryptophan (p)**	17	1239	Amide III (p)
	1575	**A, G ring-breathing (d)**		1240	**RNA peak, or asymmetric O –P–O stretching modes (d), or collagen (p)**
				1241	**Asymmetric O–P–O Stretching modes (d)**
7	1600–1800	(p)	18	1250/2	**G, C (NH_2 mode) (d)**
	1640–1680	Amide I band (p)		1254	C–N in plane stretching (p)
8	643	C–C twisting mode of tyrosine (p)		1264	Triglycerides (l)
				1265/6	Amide III, a-helix (p)
9	957	Cholesterol (l)		1453	(p)
10	977	Phenylalanine (p)	21	1465	(l)
	980	C–C stretching of beta-sheet (p),] CH bending (l)			
11	1005	Phenylalanine (p)			

[a] Abbreviations: (d) nucleic acid, (c) carbohydrate, (p) protein, (l) lipid, (G) guanine, (C) cytosine, (A) adenine, (T) thymine, (U) uracil.

Table 3.3 Provisional molecular assignments for peaks in the difference spectrum from the treatment group. Bold indicates nucleic acid-related peaks. References: [122-124]. (Reprinted with permission from ref. [104], Copyright 2015 The Royal Society of Chemistry)

The difference spectra, PC1 loadings plot and LD loadings plot shown in **Figure 3.5b** have very similar characteristics, indicating that PC1 is accounting for inter-group variation and most of the same characteristics are seen in the difference spectrum. Examining the positive features in these spectra suggests an increase in nucleic acid and protein content in the STOSE cells, as expected.

In the difference spectrum, there are negative values for lipid and some protein peaks in the treated STOSE compared to the MOSE. This contradicts previous reports where it was found that cancer cells have a higher lipid metabolism then normal cells [128,129]. This may be due to a cellular reaction to the Roscovitine treatment. The contradiction was explored in the next section by comparing the control and treated data sets of the same cell type using the same type of PCA/LDA analysis.

3.4.3. Effect of Roscovitine within groups

There was an overall decrease in the lipid peaks and some of the protein peaks in the treated groups compared to the untreated control groups. This is seen in the difference spectra (c, f in **Figure 3.6)** by the negative peaks around 960, 1070 and 1260 cm^{-1}. A strong increase is seen in the 1600 cm^{-1} peak, which is attributed to increased protein in the treated samples compared to the untreated control. The peak at 675 cm^{-1} could be attributed to nucleic acid content or could be an artifact of background subtraction as mentioned previously.

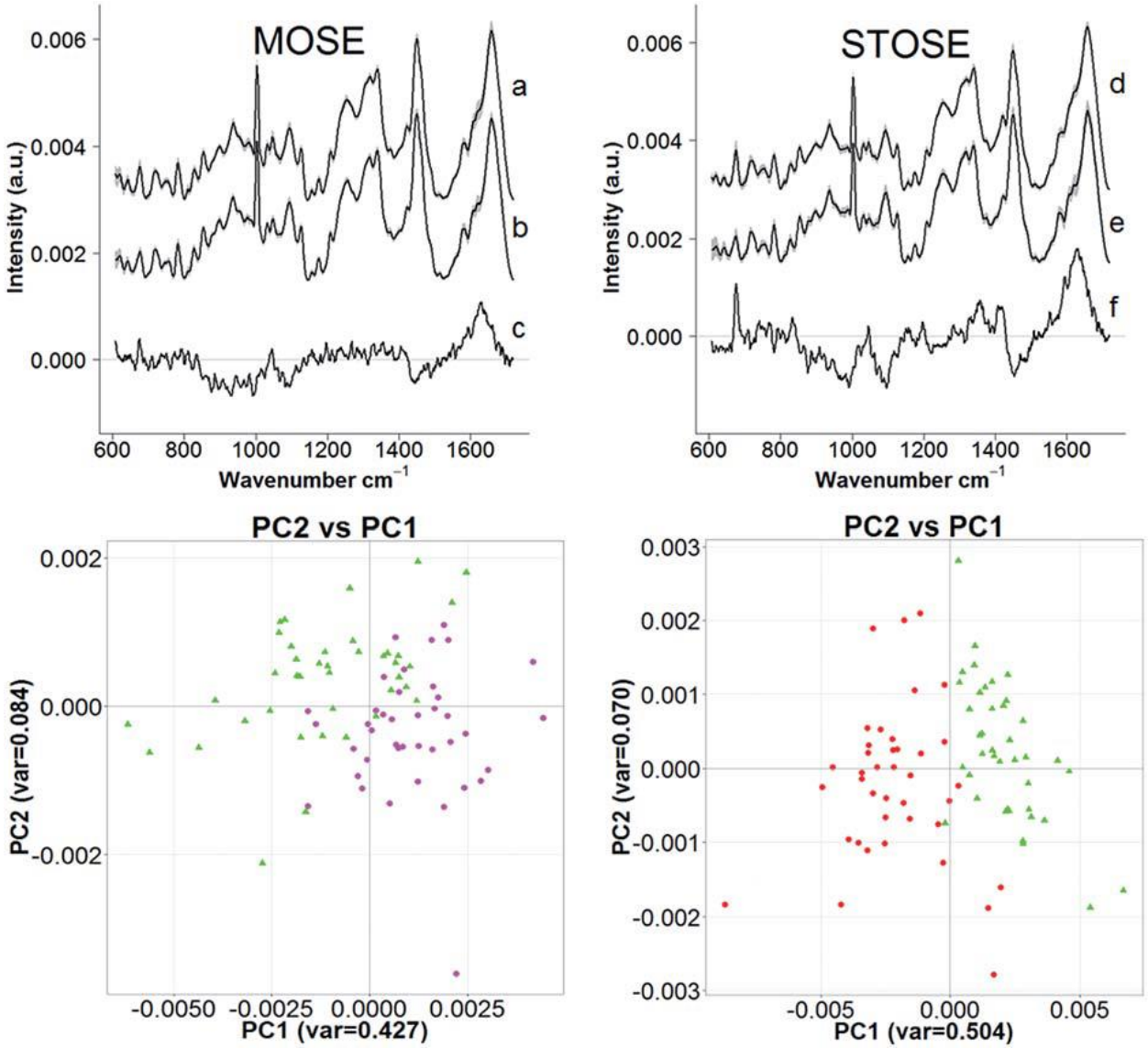

Figure 3.6 Left: Average MOSE (a) treated, (b) control and (c) difference spectra (treated-control), and PC scores plot for MOSE control and treatment spectra. Right: Average STOSE (d) control and (e) treatment, and (f) difference spectrum, and PC scores plot for STOSE control and treatment spectra. (Reprinted with permission from ref. [104], Copyright 2015 The Royal Society of Chemistry)

Comparing the PCs, there is good clustering in the PC1 vs. PC2 plots (**Figure 3.6**) for the two groups, showing that there are significant spectral differences between the treated and the untreated groups and that there is an effect of the Roscovitine treatment on eliminating cell cycle variation, although this is not fully effective. Roscovitine is known mostly to arrest cells in the G1/G0 or the G2/M phases but the majority of cells were expected to be found in the G1/G0 phase since the cells do not stay for a long time in the G2/M phase. Both observations have been reported previously [126,127], but in this study Roscovitine served the purpose of reducing the variation by eliminating the fraction of cells that are in S phase so that baseline DNA content could be compared between the MOSE and STOSE cells.

3.5. Conclusions

The PCA/LDA method can successfully discriminate between the normal MOSE and malignant STOSE cell types with an accuracy of 89% with cell cycle arrest. Further work should examine differences in the Raman spectra based on the rest of the cellular components, such as the cytoplasm. This is technically challenging with the system used in this work, as the cytoplasm is considerably thinner than the nucleus, requiring much longer and potentially damaging spectral integration times.

This initial *in vitro* study examined some of the ability of the PCA/LDA method to be able to distinguish the difference in normal and malignant ovarian cells. This had a focus on the fundamental biochemical differences and the ability to distinguish them. The multivariate technique shows a lot of promise to be a model in the future in the clinical setting to classify not only MOSE *vs.* STOSE but potentially also normal ovarian cells *vs.* cancer stem cells or ovarian cells that are sensitive or resistant to treatment. This would be advantageous in early detection of ovarian cancer, isolation of stem cells as therapeutic targets and early assessment of treatment response.

Alternate statistical techniques could also be employed to further the understanding of the biochemical composition, such as biochemical component analysis [130] or other classification methods that may be more precise, such as boosted trees [41]. Even combinations of PCA/LDA

with other techniques, such as successive projection algorithm and genetic algorithm, have shown improved results [131].

Translating this work to *in vivo* would be the next step. There are known differences that will be seen in the spectra, apart from the inability to arrest the cell cycle, and there are other components in the tissue that will also contribute to the signal. Several studies have shown the potential of this technique in a clinical setting, including endoscopic applications for colon cancer [39,132] and Barrett's esophagus [58,133]. The fiber optic probes used need to be specifically optimized for this purpose in order to minimize fluorescence background and increase the Raman signal collection. In ovarian cancer this technique could be applied during laparoscopy and will need to demonstrate high specificity for discriminating lesions and high sensitivity for their detection in order to have clinical impact and adoption.

Chapter 4
Discussion and Future Work

4.1. Summary of work to date

Diagnostic tests are important for early detection of disease. In the case of lung cancer, it has been shown that low-dose CT screening in a population with a high risk of developing the disease resulted in a 20% decrease in lung-cancer mortality [134]. This study underlined the importance of early detection. The increase in the prevented deaths from screening was correlated with the number of tumors diagnosed at stage I. Early lung cancer presents itself as a pulmonary nodule. These nodules are often false positive on screening, so that the presence of malignant tissue must be confirmed through a biopsy. Biopsy is an invasive procedure. Preforming a transthoracic needle biopsy of a pulmonary nodule presents high risk of pneumothorax, i.e. presence of air between the lung and the chest wall, causing a lung collapse, and sometimes hemorrhage [135,136]. Potentially, a less invasive procedure involving Raman spectroscopy could be considered. *In vivo* SERS imaging relies on the recognition of known markers on the tissue by using synthesized immunocomplexes with a SERS NP marker, as described in this work. Apart from just diagnosing lung cancer, antibody labeled SERS NPs have the potential to recognize a range of biomarkers present in the tissue that can lead to a more personalized diagnosis of the tumor and a more successful treatment plan [68].

Early detection of cancer is important in successful treatment of the disease. Since research has demonstrated that cancer cells originate and start to develop through pre-malignant stages, early identification and destruction of these cells can prevent the development of invasive disease. There is considerable clinical need to identify the cells undergoing these biochemical changes. Label-free Raman spectroscopy can provide a minimally- to non-invasive way to achieve this and may be used in the future as an alternative or adjuvant to histopathology. The current gold standard of histopathology is time consuming and costly with the sectioning, staining and analysis steps. The advantages of Raman spectroscopy over histopathology is that the technique, combined with multivariate statistics similar to those was described in this work, can be automated, objective and reproducible [137].

These clinical changes towards early diagnosis can be achieved using multiple techniques of Raman spectroscopy. Firstly, the intrinsic Raman spectroscopic properties of the tissue can be probed by collecting spectra containing biochemical information that can be linked to disease states. Secondly, molecular information relating to the expression levels of known disease antigens can also identify the disease. Information about a range of antigens can be collected simultaneously using molecularly targeted SERS NPs (multiplexing). This body of work aimed to develop Raman spectroscopy as a technique that can distinguish cell types by both label-free and labeled detection.

4.1.1. SERS

The first part of my work focused on the use of SERS NPs for *in vitro* cell classification. The goal was to determine the level of quantitation that could be obtained using monoclonal antibody targeted SERS NPs when binding *in vitro*. This was tested with both cells in suspension and cells adhered to a substrate. In both cases, there was a marked difference in the distribution measured by SERS NP binding compared to the expected distributions determined by fluorescent antibody staining.

The primary difficulty was identified as arising from the conjugation of the antibody to the SERS NP. This conjugation step can cause blockage of the binding site that prevents it from binding with the same affinity to the antigen site.

4.1.2. Intrinsic Raman

The second part of the body of work focused on being able to differentiate normal and malignant cells using Raman microspectroscopy *in vitro* combined with multivariate statistical analysis. These techniques have for the first time been tested in ovarian cell lines – specifically mouse ovarian surface epithelial (MOSE) cells and their spontaneously-transformed counterpart (STOSE). Although Raman microscopy and multivariate analysis have been used previously in other cell lines, this work is novel in that it shows that this can be applied in ovarian surface epithelial cells. The two groups of cells were differentiated by their Raman spectra. The

information contained in the Raman spectra also correctly showed biochemical differences between the cells. It identified that the nucleic acid content was increased in the STOSE cells compared to the MOSE.

The limitation of the technique was that the cell cycle needed to be synchronized for the PCA/LDA method to be able to classify the two types of cells with high accuracy. This synchronization was necessary to evaluate the baseline differences and identify true biochemical changes. The technique was also tested on one type of ovarian cell line. Further work needs to be done to demonstrate the use of this technique for ovarian cancer diagnosis.

4.2. Future Directions

4.2.1. Targeted SERS nanoparticles

Based on the results of our *in vitro* studies, the current formulation of the SERS NPs needs to be improved to be able to determine the extent of quantitative multiplexing that can be used. It was found that the binding affinity of the antibodies with the highest expression levels for the cell line was reduced when they are conjugated to the SERS NPs compared to the other antibody binding. A whole IgG antibody is also not the best choice for targeting. Smaller targeting moieties that are specifically developed for the antigen might increase the binding affinity. Examining the low binding affinity of the anti-EGFR conjugated SERS NPs, the antibody may be conjugated at a location that may be close to the binding site. This is very challenging to test directly. The chemistry involved the binding of the NHS ester group on the PEG linker chain to the primary amine group on the antibody. The amine group can be found at the N-terminus of the antibody and also at different points in the protein chain on the lysine amino acid. The N-terminus of the light chains in the antibody is where the binding site is formed. Conjugation to the amine group in this region can significantly affect the antigen-binding activity of the antibody. In addition, if the binding site contains a large amount of lysines, then the antigen-binding activity will be even more decreased. A further consideration is the orientation of the antibody upon conjugation to the SERS NPs. Presumably, if the antibody is conjugated near the binding site, then the binding will also be affected. To overcome these limitations of using

monoclonal antibodies, it would be advantageous to design a moiety specifically with a conjugation site that is optimized to be far from the binding site. In this way the antigen-binding activity can be maximized without the SERS NP affecting the binding significantly.

Previous work in our lab showed that anti-EGFR SERS NPs exhibited strong binding to tumor cells [33]. One difference from this work is that the antibodies used previously, namely cetuximab (a chimeric human mouse anti-EGFR monoclonal antibody) and panitumumab (a fully human antibody) are different from the mouse monoclonal antibody used here. Since these are different molecules, it is possible that the conjugation could have proceeded at a site that decreased the binding ability.

Further work by Y. Wang *et al.* [138] using a mouse monoclonal EGFR antibody and HER2 antibody concluded that binding of the multiplexed targeted SERS NPs in rat esophageal tissue can give the expected results (as compared to flow cytometry). Although the imaging results are convincing in this work, the quantitative results are not. The signal collected over a cell line derived xenograft tumor has much less separation between the binding concentrations of the different types of antibodies. Focusing on their A431 cell-line analysis, flow cytometry results of the cell line showed ~100x more anti-EFGR binding and ~40x more anti-HER2 binding than control antibody. The results from A431 cell line derived xenograft tumor found that there was a much lower ratio difference for targeted antibody binding compared to the control antibody binding of the SERS NPs. Treatment of tumor tissue showed approximately the same amount of EGFR and HER2 targeted binding (relative to the control antibody binding) over the tumor with only small specific regions showing higher targeted EGFR expression. As this was a pilot study for the endoscopic technique and only included only a single tumor, clear conclusions cannot be drawn about quantitative abilities of SERS NP binding. Other tumors did not display such discrepancies between SERS NP binding and flow cytometry. It seems like some antibodies (or clones) or cell line-derived tumors can allow for better binding of the SERS NPs than others.

There are some diseases where the detection of multiple biomarkers could be beneficial for diagnosis [139-141]. This has not become common usage in the clinic yet. Even fluorophore conjugated antibodies are not commonly used clinically for cancer diagnosis. For a diagnostic

colonoscopy for example, there have been some trials where antibody targeting has been tested [142]. Preclinical and clinical studies have shown some ability to detect dysplasia in the presence of background inflammation and to visualize inflammatory molecular targets for assessing disease severity and prognosis. However, this has yet to be adapted into routine clinical diagnosis.

Fluorescently-labeled probes are also used in fluorescent image-guided surgery [143]. They have been shown to target cancer cells [144,145], sentinel lymph nodes [146,147] and other non-cancer pathologies such as in neurological [148] and cardiovascular diseases [149]. Here is another area where SERS multiplexing could be used to localize multiple biomarkers. This would benefit the surgeon by being able to identify not only diseased tissue but also different types of healthy tissue for careful excision.

SERS NP multiplexing may be an asset in many types of disease detection for both diagnosis and surgery but the biomarkers first need to be identified to have clinical diagnostic or prognostic value. The advantage of NIR excitation used by the SERS NPs is that there is minimal tissue autofluorescence, allowing for higher levels of multiplexing, but the main advantage of NIR imaging is the ability of the signal to penetrate tissue up to centimeters [150]. It is not yet clear how deeply the NPs will penetrate into the tissue by topical application. If the nanoparticles are able to penetrate and bind to tissue deeper than just the top layer, then the collected signal may be increased due to larger number of available binding sites. This will give a better representation of the whole tissue, as multiple layers will be sampled and variation could be accounted for.

Multiplexing is a useful tool that can be potentially employed but that is still far from being implemented quantitatively in the clinical setting. The instrumentation needed for image collection of the signal is under development [35,36,33]. Most importantly, a spectrum of relevant biomarkers needs to be developed and tested with high specificity and sensitivity for the target.

4.2.2. Intrinsic Raman

4.2.2.1. In Vitro

The setup used here for signal collection was preliminary, with room for improvement to acquire more detailed cell spectra. It was selected mainly to try to preserve the most natural state of the cells to avoid any confounding effects from the environment or from the detachment process. This initial exploration with PCA/LDA analysis confirmed some biochemical changes. Further improvements include decreasing the background, increasing the number of collected spectra, and cell mapping.

The limitation of having the cells adhered to the surface is that it decreases the thickness of the cell and significantly increases the required Raman spectrum acquisition time. The signal collection was not fully confocal, leading to inefficiency in the signal collection and decreasing the signal-to-background ratio. The quartz that was used as the cell substrate is considered to have a very low Raman signal but this was still quite strong compared to the signal from the cells, which created challenges to completely remove it from the cell signal. This resulted in regions in the spectra that cannot be interpreted accurately due to interference from background. An improvement may be to use calcium fluoride coverslips, as they have been shown to provide a much lower Raman background signal.

Collection of an order-of-magnitude more spectra would allow classification of subcategories within one type of cell line. An important improvement to the workflow is to automate the system to allow more cell spectra to be collected. The 40 spectra per cell type presented in this work combined to about 3h of spectral collection and preparation time. Even though the cells were kept in an incubator that maintained them under optimal conditions, the person-hours required to manually move the stage and focus on a cell are unrealistic if hundreds of spectra are to be collected. Collecting large amounts of spectral data is important when trying to identify rare subtypes in a cell population, such as stem cells.

An optimal way to combine both of the above requirements- for faster signal collection and automation- would be to use a microfluidics setup. In this way cells can travel through the laser beam path and the Raman spectra can be collected. The advantage would be that the cells are more spherical in shape so that they have a thicker cross section than adhered cells. This would decrease the background contribution and generate a stronger Raman signal, which decreases the collection time. This also allows for the collection for the Raman spectra that represent the whole cell and that can be collected faster, since more of the cellular volume is being covered by the beam. This process could be automated so that a large number of spectra could be collected for statistical analysis. This has been tried previously for cell sorting, using optical trapping and quartz capillary channels [151] with application to analysis of erythrocytes, leukocytes, acute myeloid leukemia cells and breast tumor cells. This may also be done with the development of a Raman flow cytometry system, where Raman spectra are collected instead of fluorescence spectra as in conventional instruments [152].

For most cells to flow through the system they must be detached from their substrate. For the ovarian cells, the viability would need to be evaluated, since cells can react differently to being unable to adhere. The detachment process might affect their state, so that the collected Raman spectrum might not be an accurate representation of the true components of the cell but be confounded by the detachment procedure.

The optimal next stage from the current *in vitro* work is cell mapping. The work presented here evaluated the Raman spectra as collected mostly from the nucleus. It was determined that there is an increased amount of nucleic acid content in the malignant cells. There may be an effect due to this in the rest of the cell, so that collecting Raman spectra also from the cytoplasm is necessary to fully understand the changes occurring in the transformation of the cells from normal to malignant. Previous work in various cell lines has shown that lipid content in cancer cells is higher in the cytoplasm [128,129] compared to normal cells due to differences in lipid metabolism. From the small portion of the signal that was collected from the cytoplasm in our work, it seems like our results might contradict this conclusion (**Figure 3.5**). Although these differences could be attributed to the Roscovitine treatment, a definitive exploration of the cytoplasm content of these ovarian surface epithelial cells would help determine if the lipid

metabolism in this type of cells is atypical. If a difference from expected behavior is indeed found, then this could lead to a different use of lipid metabolism targeting cancer drugs.

4.2.2.2. In Vivo

The *in vitro* testing of the Raman capabilities established a methodology to for *in vivo* work, which is needed to establish that the technique can be used to identify malignant cells in the ovaries, recognizing that the work presented in Chapter 3 is in a very artificial setting and may differ from the findings that might be seen in real tissue samples.

The first step will be to examine tissue slices *ex vivo* that contain the mouse ovarian surface epithelial cells and compare them to slices from murine tumor formed by subcutaneous or orthotropic injection of the STOSE cells. The tissue will be more heterogeneous than the live cells grown on quartz. Different types of cells will contribute towards the collected Raman signal. It is possible that these components will drown out the differences seen in the single cell results. However, it is also possible there are other differences in the way the healthy and the malignant tissue grow that will contribute to further differences in the observed peaks.

4.2.2.3. Probes and Coherent Raman Scattering

Collecting Raman spectra in a clinical setting would require the development of a Raman fiber-optic probe to reach deep into the body. For ovarian tissue, it is not yet clear if this would be applied during a laparoscopic examination itself or to the excised tissue following biopsy. The design of the probe will depend on the sample. For example, a rigid probe would suffice for exposed tissues such as skin or brain tissue during surgery or biopsy samples. For spectral collection in body cavities such as colon or the lung a more flexible probe is required. The materials used in this probe need to be optimized, as many of them can have interfering Raman and fluorescence signals. The signal integration time needs to be minimized to be practical during diagnosis or surgery.

In order to overcome the low signal using spontaneous Raman spectroscopy, a number of techniques have been developed to enhance the signal. Stimulated interactions, where Raman scattering processes are driven coherently, generate much stronger scattered signals. The two methods that have been successfully applied in the medical field to date are stimulated Raman

scattering (SRS) and coherent anti-Stokes Raman scattering (CARS). Both methods require the use of a pumped laser beam, ω_P, and a Stokes laser beam, ω_S, where the difference in frequency ($\omega_P - \omega_S$) is tuned to match the vibrational levels of the molecule. The exploitation of such non-linear interactions provides much larger signal-to-noise ratio, thereby reducing the time needed to collect spectra from a region of tissue. Stimulated Raman is particularly exciting, since it can be used for real-time imaging instead of collecting point spectra. For example, it has been used to delineate glioma margins in brain by obtaining the ratio of the Raman signal at 2930 cm^{-1} and 2845cm^{-1}, representing the protein and lipid concentrations, respectively. Comparing SRS microscopy-analyzed biopsy samples with H&E- stained tissue gave similar accuracy results when analyzed by neuropathologists [153]. The sensitivity of SRS imaging is significantly greater than spontaneous Raman scattering and it is able to deliver background-free chemical contrast [111].

4.2.2.4. Clinical Use

Ovarian cancer has a relatively low prevalence of 40 cases per year per 100,000 women over the age of 50. If a screening test were to be developed it would need to have high sensitivity and specificity to reduce false positives and be clinically useful. It has been estimated that this would require a minimum sensitivity of 75% and a specificity that is higher at more than 99.6%. This would provide a positive predictive value of 10%, which is the threshold most commonly accepted by epidemiologists to accept a screening test [62]. This would mean there are 10 surgeries to confirm disease for every 1 case of cancer.

The technique shown in this work had a sensitivity of 92% and specificity of 85% in the case of cell-cycle synchronized cells. This has a good sensitivity but the specificity is too low to be considered as a screening test for ovarian cancer.

Another clinical application of the tool could be for use during surgical resection. Currently, epithelial ovarian cancer patients with stage IV disease have a median of 12 months progression-free survival and 29 months of overall survival. It was found that patients with a higher residual tumor after surgery have diminishing survival of both types [154] . The generally excepted belief is that optimal reduction of the tumor prolongs survival [155-157]. During surgery the surgeon

determines the boundaries of the tumor and how much to resect. The resected amount has to be balanced with the preservation of normal tissue. Raman spectroscopy can be used during surgery to find tumor margins. As mentioned in the Introduction, it has been shown to detect glioma with high enough accuracy for potential clinical adoption [41]. In the case of ovarian cancer, it would benefit the procedure for the surgeon to know with higher accuracy if the tissue has cancer cells remaining. It was shown in this work that Raman spectroscopy has the capability to detect higher DNA (nucleic acid) content found in the malignant cell line compared to the normal cell line. If this detection can be translated to tissue, then tumor margins would be better identified and localized during surgery, which would guide surgical decision making. Other work in ovarian cancer has shown that the degree of aneuploidy of tumor cells may be an indicator of patient survival [158]: higher DNA content was found in tumor cells for patients who had lower survival. In this case, preforming spectral analysis of tumor cells during surgery may also help surgeons decide the level of risk. Thus, in more sensitive areas tumor cells may be spared if they have a lower DNA content. Further research will need to be performed to determine if this is possible to determine using Raman spectroscopy and the resulting benefit.

4.3. Conclusions

Raman spectroscopy offers a wealth of biochemical information about the tissue that can contribute to diagnosis of cancer. The initial tests were carried out *in vitro* and can have potential translatability *in vivo*. Raman spectroscopy may be exploited though targeted biomarker specific SERS probes to detect specific changes in cell-surface receptors. This work is under further development to find appropriate targeting moieties that will maintain their targeting abilities upon conjugation to a SERS NP probe. Raman spectroscopy can also be used to detect changes in the tissue in a label-free manner by looking at the intrinsic spectra for biochemical changes that occur following transformation of the cells from normal to malignant.

References

1. Statistics., C.C.S.s.A.C.o.C.: Canadian cancer statistics 2013. Canadian Cancer Society, Toronto, ON (2013)

2. Austin, L.A., Osseiran, S., Evans, C.L.: Raman technologies in cancer diagnostics. Analyst (2016). doi:10.1039/C5AN01786F

3. Xie, W., Schlucker, S.: Medical applications of surface-enhanced Raman scattering. Physical chemistry chemical physics : PCCP **15**(15), 5329-5344 (2013). doi:10.1039/c3cp43858a

4. Schrader, B.: Infrared and Raman Spectroscopy: Methods and Applications. VCH, New York, NY (1995)

5. De Beer, T., Burggraeve, A., Fonteyne, M., Saerens, L., Remon, J.P., Vervaet, C.: Near infrared and Raman spectroscopy for the in-process monitoring of pharmaceutical production processes. International Journal of Pharmaceutics **417**(1–2), 32-47 (2011). doi:http://dx.doi.org/10.1016/j.ijpharm.2010.12.012

6. Paudel, A., Raijada, D., Rantanen, J.: Raman spectroscopy in pharmaceutical product design. Advanced Drug Delivery Reviews **89**, 3-20 (2015). doi:http://dx.doi.org/10.1016/j.addr.2015.04.003

7. Barhoumi, A., Zhang, D., Tam, F., Halas, N.J.: Surface-Enhanced Raman Spectroscopy of DNA. Journal of the American Chemical Society **130**(16), 5523-5529 (2008). doi:10.1021/ja800023j

8. Bell, S.E.J., Sirimuthu, N.M.S.: Surface-Enhanced Raman Spectroscopy (SERS) for Sub-Micromolar Detection of DNA/RNA Mononucleotides. Journal of the American Chemical Society **128**(49), 15580-15581 (2006). doi:10.1021/ja066263w

9. Fleischmann, M., Hendra, P.J., McQuillan, A.J.: Raman spectra of pyridine adsorbed at a silver electrode. Chemical Physics Letters **26**(2), 163-166 (1974). doi:http://dx.doi.org/10.1016/0009-2614(74)85388-1

10. Etchegoin, P.G., Le Ru, E.C.: Basic Electromagnetic Theory of SERS. In: Surface Enhanced Raman Spectroscopy. pp. 1-37. Wiley-VCH Verlag GmbH & Co. KGaA, (2010)

11. Smith, E., Dent, G.: Modern Raman Spectroscopy: A Practical Approach. J. Wiley, Hoboken, NJ (2005)

12. Institute, N.C.: Biomarker. http://www.cancer.gov/publications/dictionaries/cancer-terms?CdrID=45618.

13. Henry, N.L., Hayes, D.F.: Cancer biomarkers. Molecular Oncology **6**(2), 140-146 (2012). doi:http://dx.doi.org/10.1016/j.molonc.2012.01.010

14. Zhang, X., Soori, G., Dobleman, T.J., Xiao, G.G.: The application of monoclonal antibodies in cancer diagnosis. Expert Review of Molecular Diagnostics **14**(1), 97-106 (2014). doi:10.1586/14737159.2014.866039

15. Gonzalez, S.A., Keeffe, E.B.: Diagnosis of Hepatocellular Carcinoma: Role of Tumor Markers and Liver Biopsy. Clinics in Liver Disease **15**(2), 297-306 (2011). doi:http://dx.doi.org/10.1016/j.cld.2011.03.012

16. Taneja, S.S.: ProstaScint® Scan: Contemporary Use in Clinical Practice. Reviews in Urology **6**(Suppl 10), S19-S28 (2004).

17. Chen, Z., Fillmore, C.M., Hammerman, P.S., Kim, C.F., Wong, K.-K.: Non-small-cell lung cancers: a heterogeneous set of diseases. Nat Rev Cancer **14**(8), 535-546 (2014). doi:10.1038/nrc3775

18. Shigematsu, H., Takahashi, T., Nomura, M., Majmudar, K., Suzuki, M., Lee, H., Wistuba, I.I., Fong, K.M., Toyooka, S., Shimizu, N., Fujisawa, T., Minna, J.D., Gazdar, A.F.: Somatic Mutations of the HER2 Kinase Domain in Lung Adenocarcinomas. Cancer Research **65**(5), 1642-1646 (2005). doi:10.1158/0008-5472.can-04-4235

19. Engelman, J.A., Zejnullahu, K., Mitsudomi, T., Song, Y., Hyland, C., Park, J.O., Lindeman, N., Gale, C.-M., Zhao, X., Christensen, J., Kosaka, T., Holmes, A.J., Rogers, A.M., Cappuzzo, F., Mok, T., Lee, C., Johnson, B.E., Cantley, L.C., Jänne, P.A.: MET Amplification Leads to Gefitinib Resistance in Lung Cancer by Activating ERBB3 Signaling. Science **316**(5827), 1039-1043 (2007). doi:10.1126/science.1141478

20. Yamamoto, H., Shigematsu, H., Nomura, M., Lockwood, W.W., Sato, M., Okumura, N., Soh, J., Suzuki, M., Wistuba, I.I., Fong, K.M., Lee, H., Toyooka, S., Date, H., Lam, W.L., Minna, J.D., Gazdar, A.F.: PIK3CA Mutations and Copy Number Gains in Human Lung Cancers. Cancer Research **68**(17), 6913-6921 (2008). doi:10.1158/0008-5472.can-07-5084

21. Xing, Y., Chaudry, Q., Shen, C., Kong, K.Y., Zhau, H.E., Chung, L.W., Petros, J.A., O'Regan, R.M., Yezhelyev, M.V., Simons, J.W., Wang, M.D., Nie, S.: Bioconjugated quantum dots for multiplexed and quantitative immunohistochemistry. Nat. Protocols **2**(5), 1152-1165 (2007).

22. Ghazani, A.A., Lee, J.A., Klostranec, J., Xiang, Q., Dacosta, R.S., Wilson, B.C., Tsao, M.S., Chan, W.C.W.: High Throughput Quantification of Protein Expression of Cancer Antigens in Tissue Microarray Using Quantum Dot Nanocrystals. Nano Letters **6**(12), 2881-2886 (2006). doi:10.1021/nl062111n

23. Derfus, A.M., Chan, W.C.W., Bhatia, S.N.: Probing the Cytotoxicity of Semiconductor Quantum Dots. Nano Letters **4**(1), 11-18 (2004). doi:10.1021/nl0347334

24. Tromberg, B.J., Shah, N., Lanning, R., Cerussi, A., Espinoza, J., Pham, T., Svaasand, L., Butler, J.: Non-Invasive In Vivo Characterization of Breast Tumors Using Photon Migration Spectroscopy. Neoplasia (New York, N.Y.) **2**(1-2), 26-40 (2000).

25. Mulvaney, S.P., Musick, M.D., Keating, C.D., Natan, M.J.: Glass-Coated, Analyte-Tagged Nanoparticles: A New Tagging System Based on Detection with Surface-Enhanced Raman Scattering. Langmuir **19**(11), 4784-4790 (2003). doi:10.1021/la026706j

26. Otsuka, H., Nagasaki, Y., Kataoka, K.: PEGylated nanoparticles for biological and pharmaceutical applications. Advanced Drug Delivery Reviews **55**(3), 403-419 (2003). doi:http://dx.doi.org/10.1016/S0169-409X(02)00226-0

27. Jokerst, J.V., Miao, Z., Zavaleta, C., Cheng, Z., Gambhir, S.S.: Affibody-Functionalized Gold–Silica Nanoparticles for Raman Molecular Imaging of the Epidermal Growth Factor Receptor. Small (Weinheim an Der Bergstrasse, Germany) **7**(5), 625-633 (2011). doi:10.1002/smll.201002291

28. Herbst, R.S., Heymach, J.V., Lippman, S.M.: Lung Cancer. New England Journal of Medicine **359**(13), 1367-1380 (2008). doi:doi:10.1056/NEJMra0802714

29. Maemondo, M., Inoue, A., Kobayashi, K., Sugawara, S., Oizumi, S., Isobe, H., Gemma, A., Harada, M., Yoshizawa, H., Kinoshita, I., Fujita, Y., Okinaga, S., Hirano, H., Yoshimori, K., Harada, T., Ogura, T., Ando, M., Miyazawa, H., Tanaka, T., Saijo, Y., Hagiwara, K., Morita, S., Nukiwa, T.: Gefitinib or Chemotherapy for Non–Small-Cell Lung Cancer with Mutated EGFR. New England Journal of Medicine **362**(25), 2380-2388 (2010). doi:doi:10.1056/NEJMoa0909530

30. McVeigh, P.Z., Mallia, R.J., Veilleux, I., Wilson, B.C.: Widefield quantitative multiplex surface enhanced Raman scattering imaging in vivo. BIOMEDO **18**(4), 046011-046011 (2013). doi:10.1117/1.JBO.18.4.046011

31. Dinish, U.S., Balasundaram, G., Chang, Y.-T., Olivo, M.: Actively Targeted In Vivo Multiplex Detection of Intrinsic Cancer Biomarkers Using Biocompatible SERS Nanotags. Scientific Reports **4**, 4075 (2014). doi:10.1038/srep04075

32. Wang, Y.W., Khan, A., Som, M., Wang, D., Chen, Y., Leigh, S.Y., Meza, D., McVeigh, P.Z., Wilson, B.C., Liu, J.T.C.: Rapid ratiometric biomarker detection with topically applied SERS nanoparticles. Technology **2**(2), 118-132 (2014). doi:10.1142/S2339547814500125

33. Mallia, R.J., McVeigh, P.Z., Fisher, C.J., Veilleux, I., Wilson, B.C.: Wide-field multiplexed imaging of EGFR-targeted cancers using topical application of NIR SERS nanoprobes. Nanomedicine **10**(1), 89-101 (2014). doi:10.2217/nnm.14.80

34. Wilson, B.C., Borel, S.: Nanoparticle-Enabled Optical Endoscopy: Extending the Frontiers of Diagnosis and Treatment. In: Olivo, M., Dinish, U.S. (eds.) Frontiers in Biophotonics for Translational Medicine, vol. 3. Progress in Optical Science and Photonics, pp. 273-305. Springer Singapore, (2016)

35. Garai, E., Sensarn, S., Zavaleta, C.L., Loewke, N.O., Rogalla, S., Mandella, M.J., Felt, S.A., Friedland, S., Liu, J.T.C., Gambhir, S.S., Contag, C.H.: A Real-Time Clinical Endoscopic System for Intraluminal, Multiplexed Imaging of Surface-Enhanced Raman Scattering Nanoparticles. PLoS ONE **10**(4), e0123185 (2015). doi:10.1371/journal.pone.0123185

36. Wang, Y.W., Khan, A., Leigh, S.Y., Wang, D., Chen, Y., Meza, D., Liu, J.T.C.: Comprehensive spectral endoscopy of topically applied SERS nanoparticles in the rat esophagus. Biomedical Optics Express **5**(9), 2883-2895 (2014). doi:10.1364/BOE.5.002883

37. Votteler, M., Carvajal Berrio, D.A., Pudlas, M., Walles, H., Schenke-Layland, K.: Non-contact, Label-free Monitoring of Cells and Extracellular Matrix using Raman Spectroscopy. Journal of Visualized Experiments : JoVE(63), 3977 (2012). doi:10.3791/3977

38. Matthews, Q., Brolo, A.G., Lum, J., Duan, X., Jirasek, A.: Raman spectroscopy of single human tumour cells exposed to ionizing radiation in vitro. Physics in Medicine and Biology **56**(1), 19 (2011).

39. Molckovsky, A., Song, L.-M.W.K., Shim, M.G., Marcon, N.E., Wilson, B.C.: Diagnostic potential of near-infrared Raman spectroscopy in the colon: Differentiating adenomatous from hyperplastic polyps. Gastrointestinal Endoscopy **57**(3), 396-402 (2003). doi:http://dx.doi.org/10.1067/mge.2003.105

40. Widjaja, E., Zheng, W., Huang, Z.: Classification of colonic tissues using near-infrared Raman spectroscopy and support vector machines. International Journal of Oncology **32**(3), 653-662 (2008).

41. Jermyn, M., Mok, K., Mercier, J., Desroches, J., Pichette, J., Saint-Arnaud, K., Bernstein, L., Guiot, M.-C., Petrecca, K., Leblond, F.: Intraoperative brain cancer detection with Raman spectroscopy in humans. Science Translational Medicine **7**(274), 274ra219-274ra219 (2015). doi:10.1126/scitranslmed.aaa2384

42. Yu, C., Gestl, E., Eckert, K., Allara, D., Irudayaraj, J.: Characterization of human breast epithelial cells by confocal Raman microspectroscopy. Cancer Detection and Prevention **30**(6), 515-522 (2006). doi:http://dx.doi.org/10.1016/j.cdp.2006.10.007

43. Harvey, T.J., Hughes, C., Ward, A.D., Faria, E.C., Henderson, A., Clarke, N.W., Brown, M.D., Snook, R.D., Gardner, P.: Classification of fixed urological cells using Raman tweezers. Journal of Biophotonics **2**(1-2), 47-69 (2009). doi:10.1002/jbio.200810061

44. Chan, J.W., Taylor, D.S., Lane, S.M., Zwerdling, T., Tuscano, J., Huser, T.: Nondestructive Identification of Individual Leukemia Cells by Laser Trapping Raman Spectroscopy. Analytical Chemistry **80**(6), 2180-2187 (2008). doi:10.1021/ac7022348

45. Chan, J.W., Lieu, D.K., Huser, T., Li, R.A.: Label-Free Separation of Human Embryonic Stem Cells and Their Cardiac Derivatives Using Raman Spectroscopy. Analytical Chemistry **81**(4), 1324-1331 (2009). doi:10.1021/ac801665m

46. Downes, A., Mouras, R., Bagnaninchi, P., Elfick, A.: Raman spectroscopy and CARS microscopy of stem cells and their derivatives. Journal of Raman Spectroscopy **42**(10), 1864-1870 (2011). doi:10.1002/jrs.2975

47. Downes, A., Elfick, A.: Raman Spectroscopy and Related Techniques in Biomedicine. Sensors **10**(3), 1871 (2010).

48. Schie, I.W., Huser, T.: Methods and Applications of Raman Microspectroscopy to Single-Cell Analysis. Applied Spectroscopy **67**(8), 813-828 (2013). doi:10.1366/12-06971

49. Lodish, a., Berk, A., Zipursky, S.L., Matsudaira, P., Baltimore, D., Darnell, J.: Molecular Cell Biology, 4 ed. W. H. Freeman, New York (2000)

50. Koster, A.J., Klumperman, J.: Electron Microscopy in Cell Biology: Integrating Structure and Function. Nature Reviews Molecular Cell Biology, 4: SS6-SS9 (2003).

51. Swain, R.J., Jell, G., Stevens, M.M.: Non-invasive analysis of cell cycle dynamics in single living cells with Raman micro-spectroscopy. Journal of Cellular Biochemistry **104**(4), 1427-1438 (2008). doi:10.1002/jcb.21720

52. Matthews, Q., Jirasek, A., Lum, J., Duan, X., Brolo, A.G.: 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 (2010). doi:10.1366/000370210792080966

53. Nawaz, H., Bonnier, F., Knief, P., Howe, O., Lyng, F.M., Meade, A.D., Byrne, H.J.: Evaluation of the potential of Raman microspectroscopy for prediction of chemotherapeutic response to cisplatin in lung adenocarcinoma. Analyst **135**(12), 3070-3076 (2010). doi:10.1039/C0AN00541J

54. Draux, F., Gobinet, C., Sul?-Suso, J., Manfait, M., Jeannesson, P., Sockalingum, G.D.: Raman imaging of single living cells: probing effects of non-cytotoxic doses of an anti-cancer drug. Analyst **136**(13), 2718-2725 (2011). doi:10.1039/C0AN00998A

55. Moritz, T.J., Taylor, D.S., Krol, D.M., Fritch, J., Chan, J.W.: Detection of doxorubicin-induced apoptosis of leukemic T-lymphocytes by laser tweezers Raman spectroscopy. Biomedical Optics Express **1**(4), 1138-1147 (2010). doi:10.1364/BOE.1.001138

56. Pascut, F.C., Goh, H.T., George, V., Denning, C., Notingher, I.: Toward label-free Raman-activated cell sorting of cardiomyocytes derived from human embryonic stem cells. BIOMEDO **16**(4), 045002-045002-045004 (2011). doi:10.1117/1.3570302

57. Ghita, A., Pascut, F.C., Mather, M., Sottile, V., Notingher, I.: Cytoplasmic RNA in Undifferentiated Neural Stem Cells: A Potential Label-Free Raman Spectral Marker for Assessing the Undifferentiated Status. Analytical Chemistry **84**(7), 3155-3162 (2012). doi:10.1021/ac202994e

58. Shim, M.G., Wong Kee Song, L.-M., Marcon, N.E., Wilson, B.C.: In vivo Near-infrared Raman Spectroscopy: Demonstration of Feasibility During Clinical Gastrointestinal Endoscopy ¶. Photochemistry and Photobiology **72**(1), 146-150 (2000). doi:10.1562/0031-8655(2000)0720146IVNIRS2.0.CO2

59. Kirsch, M., Schackert, G., Salzer, R., Krafft, C.: Raman spectroscopic imaging for in vivo detection of cerebral brain metastases. Anal Bioanal Chem **398**(4), 1707-1713 (2010). doi:10.1007/s00216-010-4116-7

60. Jemal, A., Siegel, R., Ward, E., Hao, Y., Xu, J., Murray, T., Thun, M.J.: Cancer Statistics, 2008. CA: A Cancer Journal for Clinicians **58**(2), 71-96 (2008). doi:10.3322/CA.2007.0010

61. Heintz, A.P.M., Odicino, F., Maisonneuve, P., Quinn, M.A., Benedet, J.L., Creasman, W.T., Ngan, H.Y.S., Pecorelli, S., Beller, U.: 26th Annual Report on the Results of Treatment in Gynecological CancerCarcinoma of the Ovary. International Journal of Gynecology & Obstetrics **95**, S161-S192 (2006). doi:http://dx.doi.org/10.1016/S0020-7292(06)60033-7

62. Clarke-Pearson, D.L.M.D.: Screening for Ovarian Cancer. The New England Journal of Medicine **361**(2), 170-177 (2009). doi:http://dx.doi.org/10.1056/NEJMcp0901926

63. Knauer, M., Ivleva, N.P., Liu, X., Niessner, R., Haisch, C.: Surface-Enhanced Raman Scattering-Based Label-Free Microarray Readout for the Detection of Microorganisms. Analytical Chemistry **82**(7), 2766-2772 (2010). doi:10.1021/ac902696y

64. Shafer-Peltier, K.E., Haynes, C.L., Glucksberg, M.R., Van Duyne, R.P.: Toward a Glucose Biosensor Based on Surface-Enhanced Raman Scattering. Journal of the American Chemical Society **125**(2), 588-593 (2003). doi:10.1021/ja028255v

65. Ren, W., Fang, Y., Wang, E.: A Binary Functional Substrate for Enrichment and Ultrasensitive SERS Spectroscopic Detection of Folic Acid Using Graphene Oxide/Ag Nanoparticle Hybrids. ACS Nano **5**(8), 6425-6433 (2011). doi:10.1021/nn201606r

66. Ock, K., Jeon, W.I., Ganbold, E.O., Kim, M., Park, J., Seo, J.H., Cho, K., Joo, S.-W., Lee, S.Y.: Real-Time Monitoring of Glutathione-Triggered Thiopurine Anticancer Drug Release in Live Cells Investigated by Surface-Enhanced Raman Scattering. Analytical Chemistry **84**(5), 2172-2178 (2012). doi:10.1021/ac2024188

67. Vo-Dinh, T., Kasili, P., Wabuyele, M.: Nanoprobes and nanobiosensors for monitoring and imaging individual living cells. Nanomedicine: Nanotechnology, Biology and Medicine **2**(1), 22-30 (2006). doi:http://dx.doi.org/10.1016/j.nano.2005.10.012

68. Vendrell, M., Maiti, K.K., Dhaliwal, K., Chang, Y.-T.: Surface-enhanced Raman scattering in cancer detection and imaging. Trends in Biotechnology **31**(4), 249-257 (2013). doi:http://dx.doi.org/10.1016/j.tibtech.2013.01.013

69. Lu, P., Wang, J., Lin, J., Lin, J., Liu, N., Huang, Z., Li, B., Zeng, H., Chen, R.: Gold nanoaggregates for probing single-living cell based on surface-enhanced Raman spectroscopy. BIOMEDO **20**(5), 051005-051005 (2014). doi:10.1117/1.JBO.20.5.051005

70. Tang, H.-W., Yang, X.B., Kirkham, J., Smith, D.A.: Chemical Probing of Single Cancer Cells with Gold Nanoaggregates by Surface-Enhanced Raman Scattering. Applied Spectroscopy **62**(10), 1060-1069 (2008). doi:10.1366/000370208786049015

71. Tang, H.-W., Yang, X.B., Kirkham, J., Smith, D.A.: Probing Intrinsic and Extrinsic Components in Single Osteosarcoma Cells by Near-Infrared Surface-Enhanced Raman Scattering. Analytical Chemistry **79**(10), 3646-3653 (2007). doi:10.1021/ac062362g

72. Lin, D., Feng, S., Pan, J., Chen, Y., Lin, J., Chen, G., Xie, S., Zeng, H., Chen, R.: Colorectal cancer detection by gold nanoparticle based surface-enhanced Raman spectroscopy of blood serum and statistical analysis. Opt. Express **19**(14), 13565-13577 (2011). doi:10.1364/OE.19.013565

73. Wang, R.F.: Tumor antigens discovery: perspectives for cancer therapy. Molecular Medicine **3**(11), 716-731 (1997).

74. Even-Desrumeaux, K., Baty, D., Chames, P.: State of the Art in Tumor Antigen and Biomarker Discovery. Cancers **3**(2), 2554-2596 (2011). doi:10.3390/cancers3022554

75. Casiano, C.A., Mediavilla-Varela, M., Tan, E.M.: Tumor-associated Antigen Arrays for the Serological Diagnosis of Cancer. Molecular & cellular proteomics : MCP **5**(10), 1745-1759 (2006). doi:10.1074/mcp.R600010-MCP200

76. Wang, Y., Lee, K., Irudayaraj, J.: Silver Nanosphere SERS Probes for Sensitive Identification of Pathogens. The Journal of Physical Chemistry C **114**(39), 16122-16128 (2010). doi:10.1021/jp1015406

77. Li, X., Zhang, J., Xu, W., Jia, H., Wang, X., Yang, B., Zhao, B., Li, B., Ozaki, Y.: Mercaptoacetic Acid-Capped Silver Nanoparticles Colloid: Formation, Morphology, and SERS Activity. Langmuir **19**(10), 4285-4290 (2003). doi:10.1021/la0341815

78. Küstner, B., Gellner, M., Schütz, M., Schöppler, F., Marx, A., Ströbel, P., Adam, P., Schmuck, C., Schlücker, S.: SERS Labels for Red Laser Excitation: Silica-Encapsulated SAMs on Tunable Gold/Silver Nanoshells. Angewandte Chemie International Edition **48**(11), 1950-1953 (2009). doi:10.1002/anie.200804518

79. Pande, S., Ghosh, S.K., Praharaj, S., Panigrahi, S., Basu, S., Jana, S., Pal, A., Tsukuda, T., Pal, T.: Synthesis of Normal and Inverted Gold−Silver Core−Shell Architectures in β-Cyclodextrin and Their Applications in SERS. The Journal of Physical Chemistry C **111**(29), 10806-10813 (2007). doi:10.1021/jp0702393

80. Rodriguez-Lorenzo, L., Krpetic, Z., Barbosa, S., Alvarez-Puebla, R.A., Liz-Marzan, L.M., Prior, I.A., Brust, M.: Intracellular mapping with SERS-encoded gold nanostars. Integrative Biology **3**(9), 922-926 (2011). doi:10.1039/C1IB00029B

81. Yuan, H., Liu, Y., Fales, A.M., Li, Y.L., Liu, J., Vo-Dinh, T.: Quantitative Surface-Enhanced Resonant Raman Scattering Multiplexing of Biocompatible Gold Nanostars for in Vitro and ex Vivo Detection. Analytical Chemistry **85**(1), 208-212 (2013). doi:10.1021/ac302510g

82. Wu, L., Wang, Z., Zong, S., Huang, Z., Zhang, P., Cui, Y.: A SERS-based immunoassay with highly increased sensitivity using gold/silver core-shell nanorods. Biosensors and Bioelectronics **38**(1), 94-99 (2012). doi:http://dx.doi.org/10.1016/j.bios.2012.05.005

83. von Maltzahn, G., Centrone, A., Park, J.-H., Ramanathan, R., Sailor, M.J., Hatton, T.A., Bhatia, S.N.: SERS-Coded Gold Nanorods as a Multifunctional Platform for Densely Multiplexed Near-Infrared Imaging and Photothermal Heating. Advanced Materials **21**(31), 3175-3180 (2009). doi:10.1002/adma.200803464

84. Wang, T., Hu, X., Dong, S.: Surfactantless Synthesis of Multiple Shapes of Gold Nanostructures and Their Shape-Dependent SERS Spectroscopy. The Journal of Physical Chemistry B **110**(34), 16930-16936 (2006). doi:10.1021/jp062486x

85. Zhu, Z., Meng, H., Liu, W., Liu, X., Gong, J., Qiu, X., Jiang, L., Wang, D., Tang, Z.: Superstructures and SERS Properties of Gold Nanocrystals with Different Shapes. Angewandte Chemie **123**(7), 1631-1634 (2011). doi:10.1002/ange.201005493

86. Brown, L.O., Doorn, S.K.: A Controlled and Reproducible Pathway to Dye-Tagged, Encapsulated Silver Nanoparticles as Substrates for SERS Multiplexing. Langmuir **24**(6), 2277-2280 (2008). doi:10.1021/la703853e

87. McVeigh, P.Z., Mallia, R.J., Veilleux, I., Wilson, B.C.: Development of a widefield SERS imaging endoscope. In: 2012, pp. 821704-821704-821706

88. Zavaleta, C.L., Smith, B.R., Walton, I., Doering, W., Davis, G., Shojaei, B., Natan, M.J., Gambhir, S.S.: Multiplexed imaging of surface enhanced Raman scattering nanotags in living mice using noninvasive Raman spectroscopy. Proceedings of the National Academy of Sciences **106**(32), 13511-13516 (2009). doi:10.1073/pnas.0813327106

89. Sinha, L., Wang, Y., Yang, C., Khan, A., Brankov, J.G., Liu, J.T.C., Tichauer, K.M.: Quantification of the binding potential of cell-surface receptors in fresh excised specimens via dual-probe modeling of SERS nanoparticles. Scientific Reports **5**, 8582 (2015). doi:10.1038/srep08582

90. MacLaughlin, C.M., Parker, E.P.K., Walker, G.C., Wang, C.: Evaluation of SERS labeling of CD20 on CLL cells using optical microscopy and fluorescence flow cytometry. Nanomedicine: Nanotechnology, Biology and Medicine **9**(1), 55-64 (2013). doi:10.1016/j.nano.2012.04.003

91. MacLaughlin, C.M., Mullaithilaga, N., Yang, G., Ip, S.Y., Wang, C., Walker, G.C.: Surface-Enhanced Raman Scattering Dye-Labeled Au Nanoparticles for Triplexed Detection of Leukemia and Lymphoma Cells and SERS Flow Cytometry. Langmuir **29**(6), 1908-1919 (2013). doi:10.1021/la303931c

92. Wang, Y.W., Kang, S., Liu, J.T.: Multiplexed Molecular Imaging with Targeted SERS Nanoparticles for Rapid Tumor Detection. In: Biomedical Optics 2016, Fort Lauderdale, Florida, 2016/04/25 2016. OSA Technical Digest (online), p. CTh2A.4. Optical Society of America

93. Maiti, K.K., Dinish, U.S., Samanta, A., Vendrell, M., Soh, K.-S., Park, S.-J., Olivo, M., Chang, Y.-T.: Multiplex targeted in vivo cancer detection using sensitive near-infrared SERS nanotags. Nano Today **7**(2), 85-93 (2012). doi:http://dx.doi.org/10.1016/j.nantod.2012.02.008

94. Jain, P.K., Lee, K.S., El-Sayed, I.H., El-Sayed, M.A.: Calculated Absorption and Scattering Properties of Gold Nanoparticles of Different Size, Shape, and Composition: Applications in Biological Imaging and Biomedicine. The Journal of Physical Chemistry B **110**(14), 7238-7248 (2006). doi:10.1021/jp057170o

95. Leigh, S.Y., Som, M., Liu, J.T.C.: Method for Assessing the Reliability of Molecular Diagnostics Based on Multiplexed SERS-Coded Nanoparticles. PLoS ONE **8**(4), e62084 (2013). doi:10.1371/journal.pone.0062084

96. Hendriks, B.S., Klinz, S.G., Reynolds, J.G., Espelin, C.W., Gaddy, D.F., Wickham, T.J.: Impact of Tumor HER2/ERBB2 Expression Level on HER2-Targeted Liposomal Doxorubicin-Mediated Drug Delivery: Multiple Low-Affinity Interactions Lead to a Threshold Effect. Molecular Cancer Therapeutics **12**(9), 1816-1828 (2013). doi:10.1158/1535-7163.mct-13-0180

97. Trubetskoy, V.S., Torchilin, V.P., Kennel, S.J., Huang, L.: Use of N-terminal modified poly(L-lysine)-antibody conjugate as a carrier for targeted gene delivery in mouse lung endothelial cells. Bioconjugate Chemistry **3**(4), 323-327 (1992). doi:10.1021/bc00016a011

98. Vira, S., Mekhedov, E., Humphrey, G., Blank, P.S.: Fluorescent labeled antibodies - balancing functionality and degree of labeling. Analytical biochemistry **402**(2), 146-150 (2010). doi:10.1016/j.ab.2010.03.036

99. Nath, N., Godat, B., Benink, H., Urh, M.: On-bead antibody-small molecule conjugation using high-capacity magnetic beads. Journal of Immunological Methods **426**, 95-103 (2015). doi:http://dx.doi.org/10.1016/j.jim.2015.08.008

100. Schulze, H.G., Konorov, S.O., Caron, N.J., Piret, J.M., Blades, M.W., Turner, R.F.B.: Assessing Differentiation Status of Human Embryonic Stem Cells Noninvasively Using Raman Microspectroscopy. Analytical Chemistry **82**(12), 5020-5027 (2010). doi:10.1021/ac902697q

101. McCloskey, C.W., Goldberg, R.L., Carter, L.E., Gamwell, L.F., Al-Hujaily, E.M., Collins, O., Macdonald, E.A., Garson, K., Daneshmand, M., Carmona, E., Vanderhyden, B.C.: A new spontaneously transformed syngeneic model of high-grade serous ovarian cancer with a tumor-initiating cell population. Frontiers in Oncology **4** (2014). doi:10.3389/fonc.2014.00053

102. Kurman, R.J., Shih, I.-M.: The Origin and Pathogenesis of Epithelial Ovarian Cancer- a Proposed Unifying Theory. The American journal of surgical pathology **34**(3), 433-443 (2010). doi:10.1097/PAS.0b013e3181cf3d79

103. Crum, C.P., Herfs, M., Ning, G., Bijron, J.G., Howitt, B.E., Jimenez, C.A., Hanamornroongruang, S., McKeon, F.D., Xian, W.: Through the glass darkly: intraepithelial neoplasia, top-down differentiation and the road to pelvic serous cancer. The Journal of pathology **231**(4), 402-412 (2013). doi:10.1002/path.4263

104. Borel, S., Prikryl, E.A., Vuong, N.H., Jonkman, J., Vanderhyden, B., Wilson, B.C., Murugkar, S.: Discrimination of normal and malignant mouse ovarian surface epithelial cells in vitro using Raman microspectroscopy. Analytical Methods **7**(22), 9520-9528 (2015). doi:10.1039/C5AY02462E

105. Gamwell, L.F., Collins, O., Vanderhyden, B.C.: The Mouse Ovarian Surface Epithelium Contains a Population of LY6A (SCA-1) Expressing Progenitor Cells That Are Regulated by Ovulation-Associated Factors. Biology of Reproduction **87**(4), 80, 81-10 (2012). doi:10.1095/biolreprod.112.100347

106. Matthews, Q., Jirasek, A., Lum, J., Duan, X., Brolo, A.G.: 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 (2010). doi:10.1366/000370210792080966

107. Thomas, F.H., Ethier, J.-F., Shimasaki, S., Vanderhyden, B.C.: Follicle-Stimulating Hormone Regulates Oocyte Growth by Modulation of Expression of Oocyte and Granulosa Cell Factors. Endocrinology **146**(2), 941-949 (2005). doi:doi:10.1210/en.2004-0826

108. Beier, B.D., Berger, A.J.: Method for automated background subtraction from Raman spectra containing known contaminants. Analyst **134**(6), 1198-1202 (2009). doi:10.1039/B821856K

109. Notingher, I., Verrier, S., Haque, S., Polak, J.M., Hench, L.L.: Spectroscopic study of human lung epithelial cells (A549) in culture: Living cells versus dead cells. Biopolymers **72**(4), 230-240 (2003). doi:10.1002/bip.10378

110. Verrier, S., Notingher, I., Polak, J.M., Hench, L.L.: In situ monitoring of cell death using Raman microspectroscopy. Biopolymers **74**(1-2), 157-162 (2004). doi:10.1002/bip.20063

111. Freudiger, C.W., Min, W., Saar, B.G., Lu, S., Holtom, G.R., He, C., Tsai, J.C., Kang, J.X., Xie, X.S.: Label-Free Biomedical Imaging with High Sensitivity by Stimulated Raman Scattering Microscopy. Science **322**(5909), 1857-1861 (2008). doi:10.1126/science.1165758

112. Shlens, J.: A tutorial on principal component analysis, Systems Neurobiology Laboratory, University of California at San Diego. URL http://www. snl. salk. edu/~ shlens/pca. pdf (2005).

113. Cord, M., Cunningham, P., Joshi, D.: Machine Learning Techniques for Multimedia: Case Studies on Organization and Retrieval. Journal of Electronic Imaging **18**(3), 039901-039901-039902 (2009).

114. Larraona-Puy, M., Ghita, A., Zoladek, A., Perkins, W., Varma, S., Leach, I.H., Koloydenko, A.A., Williams, H., Notingher, I.: Development of Raman microspectroscopy for automated detection and imaging of basal cell carcinoma. BIOMEDO **14**(5), 054031-054031-054010 (2009). doi:10.1117/1.3251053

115. Fullwood, L.M., Clemens, G., Griffiths, D., Ashton, K., Dawson, T.P., Lea, R.W., Davis, C., Bonnier, F., Byrne, H.J., Baker, M.J.: Investigating the use of Raman and immersion Raman spectroscopy for spectral histopathology of metastatic brain cancer and primary sites of origin. Analytical Methods **6**(12), 3948-3961 (2014). doi:10.1039/C3AY42190B

116. Lima, K.s.M.G., Gajjar, K., Valasoulis, G., Nasioutziki, M., Kyrgiou, M., Karakitsos, P., Paraskevaidis, E., Martin Hirsch, P.L., Martin, F.L.: Classification of cervical cytology for human papilloma virus (HPV) infection using biospectroscopy and variable selection techniques. Analytical Methods **6**(24), 9643-9652 (2014). doi:10.1039/C4AY01736F

117. Pichardo-Molina, J.L., Frausto-Reyes, C., Barbosa-García, O., Huerta-Franco, R., González-Trujillo, J.L., Ramírez-Alvarado, C.A., Gutiérrez-Juárez, G., Medina-Gutiérrez, C.: Raman spectroscopy and multivariate analysis of serum samples from breast cancer patients. Lasers Med Sci **22**(4), 229-236 (2007). doi:10.1007/s10103-006-0432-8

118. Crow, P., Molckovsky, A., Stone, N., Uff, J., Wilson, B., WongKeeSong, L.M.: Assessment of fiberoptic near-infrared raman spectroscopy for diagnosis of bladder and prostate cancer. Urology **65**(6), 1126-1130 (2005). doi:http://dx.doi.org/10.1016/j.urology.2004.12.058

119. Team, R.C.: R: A language and environment for statistical computing. R Foundation for Statistical Computing. http://www.R-project.org/ (2013)

120. Venables, W.N., Ripley, B.D.: Modern Applied Statistics with S, 4th ed. Springer, New York (2002)

121. Cunningham, P.: Dimension Reduction. In: Cord, M., Cunningham, P. (eds.) Machine Learning Techniques for Multimedia. Cognitive Technologies, pp. 91-112. Springer Berlin Heidelberg, (2008)

122. Movasaghi, Z., Rehman, S., Rehman, I.U.: Raman Spectroscopy of Biological Tissues. Applied Spectroscopy Reviews **42**(5), 493-541 (2007). doi:10.1080/05704920701551530

123. Short, K.W., Carpenter, S., Freyer, J.P., Mourant, J.R.: Raman Spectroscopy Detects Biochemical Changes Due to Proliferation in Mammalian Cell Cultures. Biophysical Journal **88**(6), 4274-4288 (2005). doi:10.1529/biophysj.103.038604

124. Stone, N., Kendall, C., Smith, J., Crow, P., Barr, H.: Raman spectroscopy for identification of epithelial cancers. Faraday Discussions **126**(0), 141-157 (2004). doi:10.1039/B304992B

125. Delfino, I., Perna, G., Lasalvia, M., Capozzi, V., Manti, L., Camerlingo, C., Lepore, M.: Visible micro-Raman spectroscopy of single human mammary epithelial cells exposed to x-ray radiation. BIOMEDO **20**(3), 035003-035003 (2015). doi:10.1117/1.JBO.20.3.035003

126. Wu, P.-C., Tai, M.-H., Hu, D.-N., Lai, C.-H., Chen, Y.-H., Wu, Y.-C., Tsai, C.-L., Shin, S.-J., Kuo, H.-K.: Cyclin-dependent Kinase Inhibitor Roscovitine Induces Cell Cycle Arrest

and Apoptosis in Rabbit Retinal Pigment Epithelial Cells. Journal of Ocular Pharmacology and Therapeutics **24**(1), 25-33 (2008). doi:10.1089/jop.2007.0044

127. Cho, S.R., Ock, S.A., Yoo, J.G., Mohana kumar, B., Choe, S.Y., Rho, G.J.: Effects of Confluent, Roscovitine Treatment and Serum Starvation on the Cell-cycle Synchronization of Bovine Foetal Fibroblasts. Reproduction in Domestic Animals **40**(2), 171-176 (2005). doi:10.1111/j.1439-0531.2005.00577.x

128. Daniëls, V.W., Smans, K., Royaux, I., Chypre, M., Swinnen, J.V., Zaidi, N.: Cancer Cells Differentially Activate and Thrive on De Novo Lipid Synthesis Pathways in a Low-Lipid Environment. PLoS ONE **9**(9), e106913 (2014). doi:10.1371/journal.pone.0106913

129. Mukherjee, A., Wu, J., Barbour, S., Fang, X.: Lysophosphatidic Acid Activates Lipogenic Pathways and de Novo Lipid Synthesis in Ovarian Cancer Cells. Journal of Biological Chemistry **287**(30), 24990-25000 (2012). doi:10.1074/jbc.M112.340083

130. Ong, Y.H., Lim, M., Liu, Q.: Comparison of principal component analysis and biochemical component analysis in Raman spectroscopy for the discrimination of apoptosis and necrosis in K562 leukemia cells. Opt. Express **20**(20), 22158-22171 (2012). doi:10.1364/OE.20.022158

131. Theophilou, G., Lima, K.M.G., Martin-Hirsch, P.L., Stringfellow, H.F., Martin, F.L.: ATR-FTIR spectroscopy coupled with chemometric analysis discriminates normal, borderline and malignant ovarian tissue: classifying subtypes of human cancer. Analyst **141**(2), 585-594 (2016). doi:10.1039/C5AN00939A

132. Andrade, P.O., Bitar, R.A., Yassoyama, K., Martinho, H., Santo, A.M.E., Bruno, P.M., Martin, A.A.: Study of normal colorectal tissue by FT-Raman spectroscopy. Anal Bioanal Chem **387**(5), 1643-1648 (2007). doi:10.1007/s00216-006-0819-1

133. Kendall, C., Stone, N., Shepherd, N., Geboes, K., Warren, B., Bennett, R., Barr, H.: Raman spectroscopy, a potential tool for the objective identification and classification of neoplasia in Barrett's oesophagus. The Journal of Pathology **200**(5), 602-609 (2003). doi:10.1002/path.1376

134. Kovalchik, S.A., Tammemagi, M., Berg, C.D., Caporaso, N.E., Riley, T.L., Korch, M., Silvestri, G.A., Chaturvedi, A.K., Katki, H.A.: Targeting of Low-Dose CT Screening According to the Risk of Lung-Cancer Death. New England Journal of Medicine **369**(3), 245-254 (2013). doi:doi:10.1056/NEJMoa1301851

135. Wiener, R.S., Schwartz, L.M., Woloshin, S., Welch, H.G.: Population-Based Risk for Complications After Transthoracic Needle Lung Biopsy of a Pulmonary Nodule: An Analysis of Discharge Records. Annals of Internal Medicine **155**(3), 137-144 (2011). doi:10.7326/0003-4819-155-3-201108020-00003

136. Saji, H., Nakamura, H., Tsuchida, T., Tsuboi, M., Kawate, N., Konaka, C., Kato, H.: The incidence and the risk of pneumothorax and chest tube placement after percutaneous ct-guided lung biopsy*: The angle of the needle trajectory is a novel predictor. Chest **121**(5), 1521-1526 (2002). doi:10.1378/chest.121.5.1521

137. Kallaway, C., Almond, L.M., Barr, H., Wood, J., Hutchings, J., Kendall, C., Stone, N.: Advances in the clinical application of Raman spectroscopy for cancer diagnostics. Photodiagnosis and Photodynamic Therapy **10**(3), 207-219 (2013). doi:http://dx.doi.org/10.1016/j.pdpdt.2013.01.008

138. Wang, Y.W., Kang, S., Khan, A., Bao, P.Q., Liu, J.T.C.: In vivo multiplexed molecular imaging of esophageal cancer via spectral endoscopy of topically applied SERS nanoparticles. Biomedical Optics Express **6**(10), 3714-3723 (2015). doi:10.1364/BOE.6.003714

139. Wang, T.J., Gona, P., Larson, M.G., Tofler, G.H., Levy, D., Newton-Cheh, C., Jacques, P.F., Rifai, N., Selhub, J., Robins, S.J., Benjamin, E.J., D'Agostino, R.B., Vasan, R.S.: Multiple Biomarkers for the Prediction of First Major Cardiovascular Events and Death. New England Journal of Medicine **355**(25), 2631-2639 (2006). doi:doi:10.1056/NEJMoa055373

140. Shariat, S.F., Karakiewicz, P.I., Ashfaq, R., Lerner, S.P., Palapattu, G.S., Cote, R.J., Sagalowsky, A.I., Lotan, Y.: Multiple biomarkers improve prediction of bladder cancer recurrence and mortality in patients undergoing cystectomy. Cancer **112**(2), 315-325 (2008). doi:10.1002/cncr.23162

141. Dhanasekaran, S.M., Barrette, T.R., Ghosh, D., Shah, R., Varambally, S., Kurachi, K., Pienta, K.J., Rubin, M.A., Chinnaiyan, A.M.: Delineation of prognostic biomarkers in prostate cancer. Nature **412**(6849), 822-826 (2001). doi:http://www.nature.com/nature/journal/v412/n6849/suppinfo/412822a0_S1.html

142. Kim, S.-Y., Myung, S.-J.: Optical Molecular Imaging for Diagnosing Intestinal Diseases. Clinical Endoscopy **46**(6), 620-626 (2013). doi:10.5946/ce.2013.46.6.620

143. Vahrmeijer, A.L., Hutteman, M., van der Vorst, J.R., van de Velde, C.J.H., Frangioni, J.V.: Image-guided cancer surgery using near-infrared fluorescence. Nat Rev Clin Oncol **10**(9), 507-518 (2013). doi:10.1038/nrclinonc.2013.123

144. Weissleder, R., Tung, C.-H., Mahmood, U., Bogdanov, A.: In vivo imaging of tumors with protease-activated near-infrared fluorescent probes. Nat Biotech **17**(4), 375-378 (1999).

145. Olson, E.S., Jiang, T., Aguilera, T.A., Nguyen, Q.T., Ellies, L.G., Scadeng, M., Tsien, R.Y.: Activatable cell penetrating peptides linked to nanoparticles as dual probes for in vivo fluorescence and MR imaging of proteases. Proceedings of the National Academy of Sciences of the United States of America **107**(9), 4311-4316 (2010). doi:10.1073/pnas.0910283107

146. Hutteman, M., Van Der Vorst, J.R., Gaarenstroom, K.N., Peters, A.A.W., Mieog, J.S.D., Schaafsma, B.E., Löwik, C.W.G.M., Frangioni, J.V., Van De Velde, C.J.H., Vahrmeijer, A.L.: Optimization of Near-Infrared Fluorescent Sentinel Lymph Node Mapping for Vulvar Cancer. American journal of obstetrics and gynecology **206**(1), 89.e81-89.e85 (2012). doi:10.1016/j.ajog.2011.07.039

147. Brouwer, O.R., Klop, W.M.C., Buckle, T., Vermeeren, L., van den Brekel, M.W.M., Balm, A.J.M., Nieweg, O.E., Valdés Olmos, R.A., van Leeuwen, F.W.B.: Feasibility of Sentinel Node Biopsy in Head and Neck Melanoma Using a Hybrid Radioactive and Fluorescent Tracer. Annals of Surgical Oncology **19**(6), 1988-1994 (2012). doi:10.1245/s10434-011-2180-7

148. Hyde, D., de Kleine, R., MacLaurin, S.A., Miller, E., Brooks, D.H., Krucker, T., Ntziachristos, V.: Hybrid FMT–CT imaging of amyloid-β plaques in a murine Alzheimer's disease model. NeuroImage **44**(4), 1304-1311 (2009). doi:http://dx.doi.org/10.1016/j.neuroimage.2008.10.038

149. Deguchi, J.-o., Aikawa, M., Tung, C.-H., Aikawa, E., Kim, D.-E., Ntziachristos, V., Weissleder, R., Libby, P.: Inflammation in Atherosclerosis: Visualizing Matrix Metalloproteinase Action in Macrophages In Vivo. Circulation **114**(1), 55-62 (2006). doi:10.1161/circulationaha.106.619056

150. Chance, B.: Near-Infrared Images Using Continuous, Phase-Modulated, and Pulsed Light with Quantitation of Blood and Blood Oxygenationa. Annals of the New York Academy of Sciences **838**(1), 29-45 (1998). doi:10.1111/j.1749-6632.1998.tb08185.x

151. Dochow, S., Krafft, C., Neugebauer, U., Bocklitz, T., Henkel, T., Mayer, G., Albert, J., Popp, J.: Tumour cell identification by means of Raman spectroscopy in combination with optical traps and microfluidic environments. Lab on a Chip **11**(8), 1484-1490 (2011). doi:10.1039/C0LC00612B

152. Watson, D.A., Brown, L.O., Gaskill, D.F., Naivar, M., Graves, S.W., Doorn, S.K., Nolan, J.P.: A flow cytometer for the measurement of Raman spectra. Cytometry Part A **73A**(2), 119-128 (2008). doi:10.1002/cyto.a.20520

153. Zhang, R.R., Kuo, J.S.: Detection of Human Brain Tumor Infiltration With Quantitative Stimulated Raman Scattering Microscopy. Neurosurgery **78**(4) (2016).

154. Winter, W.E., Maxwell, G.L., Tian, C., Sundborg, M.J., Rose, G.S., Rose, P.G., Rubin, S.C., Muggia, F., McGuire, W.P.: Tumor Residual After Surgical Cytoreduction in Prediction of Clinical Outcome in Stage IV Epithelial Ovarian Cancer: A Gynecologic Oncology Group Study. Journal of Clinical Oncology **26**(1), 83-89 (2008).

155. Hacker, N.F., Berek, J.S., Lagasse, L.D., Nieberg, R.K., Elashoff, R.M.: Primary Cytoreductive Surgery for Epithelial Ovarian Cancer. Obstetrics & Gynecology **61**(4) (1983).

156. Bertelsen, K.: Tumor reduction surgery and long-term survival in advanced ovarian cancer: A DACOVA study. Gynecologic Oncology **38**(2), 203-209. doi:10.1016/0090-8258(90)90042-J

157. Delgado, G., Oram, D.H., Petrilli, E.S.: Stage III epithelial ovarian cancer: The role of maximal surgical reduction. Gynecologic Oncology **18**(3), 293-298 (1984). doi:http://dx.doi.org/10.1016/0090-8258(84)90040-4

158. Friedlander, M.L., Hedley, D.W., Taylor, I.W., Russell, P., Coates, A.S., Tattersall, M.H.N.: Influence of Cellular DNA Content on Survival in Advanced Ovarian Cancer. Cancer Research **44**(1), 397-400 (1984).

Copyright Acknowledgements

Parts of Chapter 3 are published in Analytical Methods, 2015, 7 (22), pp 9520–9528
DOI: 10.1039/C5AY02462E