

University of Nottingham
UK | CHINA | MALAYSIA

Tissue characterisation by Raman spectroscopies

by **Manon Fanny BOURBOUSSON**

Thesis submitted to the University of Nottingham
for the degree of Doctor of Philosophy

supervised by
Prof Ioan NOTINGHER
David BALDWIN

School of Physics and Astronomy
April 2019

Abstract

Cancer has a major impact on society in the United Kingdom and across the world. Lung cancer is the most common cancer worldwide. Advances in lung cancer diagnosis are essential to reduce the mortality burden. The current histopathology methods for the diagnosis of human lung tissues require highly specialised staff, are labour-intensive, and expensive. They could benefit from the development of new imaging techniques. In this work, we demonstrate the ability to use Raman spectroscopy as a non-destructive method for the analysis of the molecular composition of human lung tissues *ex vivo*. For different aspects of cancer research, non-destructive techniques are appreciated because they do not induce damage to the sample measured making them a possible additional technique in the cancer diagnosis routine. Raman spectroscopy is a chemical analysis method that detects molecular differences in biological tissue samples. It is a fast, accurate, and non-destructive approach well-suited for cancer detection. It has already been established that Raman spectroscopy may provide an improvement in diagnostic accuracy while at the same time reducing subjectivity in cancer studies compared to the standard techniques currently used in hospitals.

The thesis starts by describing the experimental instrumentations developed to achieve Raman mapping measurements of biological tissues. The study began with the investigation of lung tissues by Raman spectroscopy mapping. The instrument we designed was assembled at Nottingham City Hospital. The system was aligned and calibrated with standard materials presenting established Raman shift frequencies. Then it was optimised minimizing the Raman shift error, and fully characterised in terms of spatial and depth resolution, and stability.

The instrument allowed the collection of Raman spectral maps of fresh lung tissues, including cancerous samples, with high-quality spectra recorded in mapping

mode with 20 μm step size and an acquisition time of 1 s per spectrum. A detailed spectral analysis was achieved showing the ability to extract molecular information from 18 freshly excised lung tissues including malignant and non-malignant tissues. We present the molecular markers detected in the Raman signal that differentiate tumour tissue from healthy tissue and from necrotic tissue. The precision of our instrument allowed us to detect bands never reported before in the literature.

Then, our aim was to implement modifications to the experimental instrument design to limit the collection of tissue autofluorescence in the Raman spectra. Because the light backscattered from a sample is likely to exhibit directional dependency, we designed a new instrument allowing us to select the direction of collection of the signal, known as a polarised Raman micro-spectrometer. We performed the steps of alignment, calibration, and optimisation, by taking into account the polarisation dependence.

The performance of this instrument was firstly tested by investigating tissue samples with a significantly lower fluorescence level than lung tissue. 15 skin samples from 7 patients were investigated. The calculation of the average depolarisation ratio coefficient revealed molecular vibrations specificities of connective tissue in different local environments. Additionally, principal component analysis (PCA) was performed emphasising the variation of biomolecules orientations in connective tissue for various samples. Finally, we present a new application of the use of polarised Raman spectroscopy based on varying the polarisation direction for the collection of a reduced fluorescent background in the case of lung tissues. The results from both spectroscopy techniques developed have validated the potential of Raman spectroscopy in tissue imaging for label-free monitoring of biomolecular variations in tissues and their future clinical translation in cancer research.

Acknowledgement

I would like to express my sincere gratitude to my supervisor, Professor Ioan NOTINGHER to give me the opportunity to be a member of the Biophotonics group. His guidance, patience and weekly advice were invaluable to the completion of this work. I would like to thank David BALDWIN for the feedback and encouragement throughout the thesis.

I wish to acknowledge the Nottingham Health Science Biobank (NHSB) for their involvement in providing tissues and the histopathology team for detailed microscopic examination of the tissue samples.
I would also like to thank all my colleagues for their contribution. Particularly Dr Kenny KONG for sharing computational knowledge, Dr Zhiyu LIAO for exchanging scientific conversations, Dr Faris SINJAB for his expertise in the laboratory and Radu who never minds bother and who shared precious help for the skin project.

I am particularly grateful to have spent time with wonderful women during these years. Thanks to my office mates Manal and Somaya for their never-ending support and friendship. Thanks to my friends Naja and Marlene for their positive energy and great moments together.
Lastly, thanks to my best friend Pierre for making my time in the UK happier.

A special thanks to the french crew: Anna, Camille, Cecile, Joana and Perrette for their precious love. And finally, a warm thank you to moje kochanie Łukasz, my everyday sunshine.

To my father, to my family.
To the ones missing.

To my mom Nicole, Louise, Annonciade BOURBOUSSON who always believed in me, for the twenty-eight years of love, support and optimism.

Abbreviations and acronyms

Abbreviation	Stands for
2D	two-dimensional
3D	three-dimensional
a.u.	arbitrary unit
BCC	Basal cell carcinoma
BMB	1,4 bis (2-methylstyryl) benzene
CaF_2	Calcium fluoride
CCD	Charged-coupled Device
DNA	Deoxyribonucleic acid
EGFR	Epidermal growth factor
FTIR	Fourier-transform infrared spectroscopy
FWHM	Full Width at Half Maximum
H&E	Haematoxylin and Eosin
InGaAs	Indium Gallium Arsenide
IR	Infrared
LabVIEW	Laboratory Virtual Instrument Engineering Workbench
laser	Light Amplification by Stimulated Emission of Radiation
NA	Numerical aperture
NOS	Not otherwise specified
NIR	Near-infrared
PBS	Phosphate-buffered Saline
PCA	Principal component analysis

Abbreviation	Stands for
pH	Power of hydrogen
PpIX	Protoporphyrin IX
PRS	Polarised Raman Spectroscopy
PS	Performance status
RS	Raman spectroscopy
RNA	Ribonucleic acid
SCC	Squamous cell carcinoma
SCLC	Small cell lung carcinoma
SNR	Signal-to-noise ratio
SVD	Singular Value Decomposition
TNM	tumour-node-metastasis
TEM	Transverse electromagnetic
UK	United Kingdom
UV	Ultraviolet

Mai coume as fa toun conte?
Ai embrassa l'aubre de la paciènci. E lou cèu a tira tóuti si ridèu de nivo. Tout camin camino...

— Proverbes provencaux

Contents

1 Introduction **1**

1.1 This thesis 1

1.2 Overview 3

1.2.1 Lung cancer 4

1.2.2 Skin cancer 7

1.2.3 Screening and diagnosis tests 9

1.2.4 Limitations of the standard diagnosis methods 13

1.3 Raman spectroscopy (RS) 16

1.4 Polarised Raman spectroscopy (PRS) 20

1.5 Literature review 22

1.6 Aims 27

2 Development of Raman spectrometer instruments for RS and PRS **29**

2.1 Description of the Raman micro spectrometer 29

2.1.1 Instrumentation 30

2.1.2 Calibration 34

2.2 Description of the Polarised Raman spectrometer 44

2.2.1 Optical components 44

2.2.2 Calibration 46

2.3 Software and spectral analysis 48

3 Raman mapping studies of fresh human lung tissues **54**

3.1 Introduction 54

3.2 Material and Methods 55

Tissue collection 56

Sample handling and preparation 56

Ex-vivo Raman spectroscopy 58

Data processing 59
3.3 Results 60
Measurement parameters 60
Shrinkage effects 63
Group A: tissues with laser damage 67
Group B: tissues with strong autofluorescence . . . 68
Group C: tissues with Raman spectral information . 70
Comparison normal vs tumour tissues 75
Discussion 79
Perspectives 81

4 Characterisation of human tissue samples by polarised Raman spectroscopy 84

4.1 Introduction 84
4.2 Material and Methods 86
Sample preparation 86
Instrumentation 87
Data acquisition 87
Data processing 87
4.3 Results 88
Characterisation of skin dermis 89
Depolarisation ratio 95
Dependency of the Raman bands on the polarisation angle 98
Polarised Raman scanning at 0°, 40° and 90° . . . 100
Principal component analysis (PCA) 105
Discussion 109
4.4 Application on fluorescent tissues 110
Introduction 110
Material and methods 111
Lung tissues 111
Other tissues 112
Discussion 114

5 Conclusion **117**

5.1 General conclusions 117

5.2 Future directions 119

References **122**

1 Introduction

This chapter introduces the main objectives of the thesis. It reviews the current clinical diagnosis methods for the detection of lung and skin cancer. It provides a theoretical background on Raman spectroscopy and polarised Raman spectroscopy. It also aims to establish a link between the state of the art of spectroscopy methods for skin and lung diagnosis and our motivations.

1.1 This thesis

In the 21st century, noncommunicable diseases are the major cause of death in every country of the world [1, 2]. Among them, cancer is responsible for the largest number of deaths. The increasing incidence of cancer and mortality worldwide is, therefore, a public health priority. Prevention, early detection and effective treatment have become increasingly important to contribute to the decline of cancer death rates. Cancer detection, linked to morphological and functional variations in diseased tissues, is mostly performed by histopathology as the gold standard. Currently, in health-care institutions and for most cancer, the tissue samples are visually assessed under light microscopy. This method is prone to subjectivity error leading to diagnosis inconsistency between pathologists which can impact the patient treatment [3]. Non-invasive techniques with high sensitivity to molecular changes in tissue could supplement the human evaluation. Optical spectroscopic techniques such as fluorescence, reflectance and Raman spectroscopy are sensitive to structural molecular differences and thus meet the needs mentioned making them ideal to be used as additional methods to

histopathology. Raman spectroscopy is especially suitable for this application, and the implementation of this spectroscopy technique for tissue characterisation will be extensively discussed in this thesis. This vibrational spectroscopy consists in irradiating a sample with laser incident light which lead to the emission of Raman photons collected by a spectrometer. The Raman scattering process is the shift in frequency of the inelastically scattered radiation. The loss or gain in energy happening during the scattering process is reflected in the observed bands. The bands are assigned to specific molecular vibrational transitions providing information on the chemical composition of the sample and on the conformation of the molecules. Raman spectroscopy has been proven to be successful in a large number of studies, reviewed later, but there remain some challenges in its implementation as a clinical tool for diagnosis. There are quite a few reasons for this, such as, for some tissues the laser-induced fluorescence emission can overshadow the Raman bands making the Raman features difficult to isolate. Besides, the acquisition time influences the spectral quality and needs to be compatible with biological tissue studies because it has an impact on the laser-induced damage on tissues. The development of lab-built Raman spectrometers in combination with the great technological advances over the last decades can tackle the challenges mentioned. Indeed, their design can be optimised for specific applications, allowing the selection of the suitable adjustment for each individual component. In this work, we propose to explore the collection of Raman scattered spectra from biological samples, mainly from lung and skin tissues, induced by near-infrared laser excitation at 785 nm wavelength. The spectra collected are used as an additional diagnosis method to histopathology for cancer and pre-cancer states of human tissues. The project presented is part of a multidisciplinary research area depicting the bridge between physics and medicine and illustrating the collaboration between two fields towards a common objective.

Outline

Chapter 2 explains the instrumentation design of the lab-built Raman micro-spectrometer. It details the main components and their characteristics. Addition-

ally, we present the experimental setup developed for performing polarisation Raman microscopy. The performance characteristics of both devices are also discussed.

Chapter 3 describes the work of the development of the Raman spectrometer as a method for *ex vivo* measurement of lung tissues. The objective of this chapter is to understand molecular variations in lung tissues for improving the accuracy of tumour subtyping. Spectra were recorded for lung squamous cell carcinoma, adenocarcinoma and healthy lung tissue. The results show that the spectral differences observed allowed the differentiation of lung tumour subtypes as well as necrotic tissues.

In **Chapter 4**, we discuss the use of polarised Raman spectroscopy for tissue characterisation with depolarisation ratio analysis of typical Raman bands, band dependency with the angle of collection and principal component analysis (PCA). We also present polarised Raman spectroscopy measurements of highly fluorescent tissues and illustrate how it could help for fluorescence rejection taking advantages of the different directions of polarisation of Raman scattering and fluorescence when excited with polarised light.

In **Chapter 5**, a general conclusion from the work described for skin and lung tissues is established and the prospects of polarised Raman spectroscopy for clinical use are mentioned.

1.2 Overview

Most of the work presented in this thesis concerns lung cancer and skin cancer. These two cancer types are very common since globally, lung cancer is the leading cause of cancer death for both genders while skin cancer incidence keeps increasing over the years. It emphasises the importance of our study in the context of a major healthcare priority.

1.2.1 Lung cancer

The lungs are organs of the respiratory system situated on both sides of the chest (Figure 1.1).

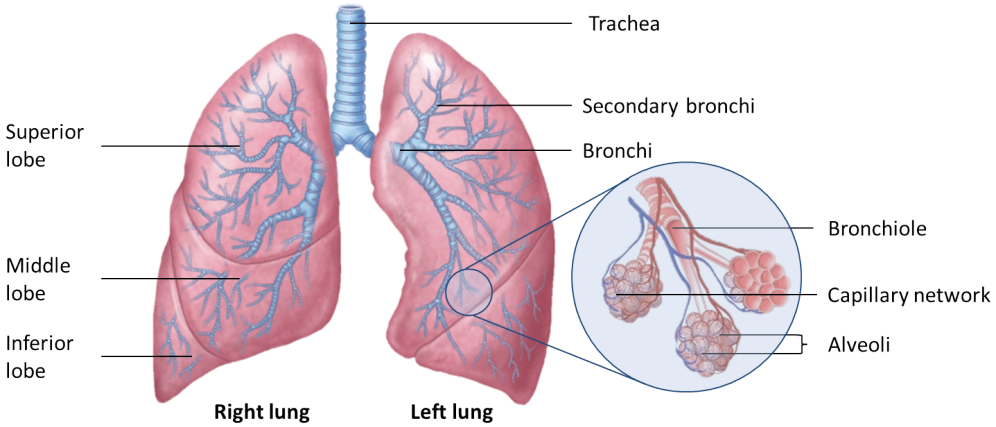


Figure 1.1: Schematic representation of the human lung anatomy.

Each lung is divided into lobes; the right lung comprises three lobes, while the left lung is divided into two lobes. The partition between the lobes insures protection from mechanical damage as well as prevention of spreading of an infection. The connection between the lungs and the atmospheric air is made by branching tubes such as bronchioles, bronchi, and trachea. The trachea connects the mouth to the lungs. The air inhaled travels through the bronchi which then divides into smaller branches called bronchioles. The bronchioles ensure the transport of the air to the end of the respiratory tree: the lung alveoli. The alveoli walls contain capillaries where oxygen and carbon dioxide are exchanged [4]. Various lung conditions are documented mainly caused by allergies, infections, air pollution and smoking but we will exclusively, in this study, describe and investigate lung cancer.

Epidemiology

Lung cancer is the deadliest cancer in the United Kingdom (UK), for both females and males. In 2016, around 35,000 people died because of lung cancer, accounting for 20 % of all cancer deaths. Concerning lung cancer mortality, the UK ranks 4th highest in Europe for women and 21st for men, with mortality in Western

Europe higher than for Central, Eastern, Southern Europe and Eastern Asia [5]. The most important cause of lung cancer is tobacco smoking. Carcinogens in cigarettes promote genetic changes that disrupt the normal functioning of cells including the development of unchecked growth. Although 85 % of lung cancer cases are attributed to tobacco, there are several other well-established causes and associations including radon exposure, family history, human immunodeficiency virus along with industrial exposure such as asbestos, arsenic or hydrocarbons. Around three-quarters of patients with lung cancer have late-stage disease, suggesting that little can be done to cure them. It is thus of paramount importance that lung cancer is diagnosed as early as possible. Even though there have been improvements in diagnosis and treatment, the 5-year survival rate of lung cancer is extremely poor, 13 % in England and 15-17 % in other similar countries [6]. Additionally, to reduce the mortality rate, it is necessary to enhance prevention through smoking cessation, and to develop better treatments. For example, new treatments, targeted to a specific subtype of the tumour, may improve outcomes in later stage disease [7].

Pathogenesis

Lung cancers are classified into two main types: non-small cell lung cancer and small-cell lung cancer. Within the first type, which covers 90 % of cases, there are three different subtypes: squamous cell carcinoma, large cell carcinoma and adenocarcinoma. Then, about 10 % of lung cancers are small cell cancer, including small-cell lung cancer and small-cell carcinoma. Squamous cell carcinoma begins in squamous cells situated in the central part of the lung near the main, lobar and segmental bronchi. They are in almost all the cases associated with smokers, they can cause obstruction in the airway leading to respiratory problems up to the collapse of a lung. Adenocarcinoma begins from cells that line in the alveoli usually found in the outer part of the lungs in the small peripheral airways. It is a cancer type common in non-smokers and women. Large cell carcinoma begins from various large cell types from any part of the lung. Small cell carcinoma begins in the bronchi near the centre of the chest. Most tumours arise from the lung epithelium in the basal layer where the cells are capable of differentiation into cells of different function and structure. Once the cancer cells have developed,

they can spread in the lymphatic vessels system. Cell growth in lymph nodes presents a high risk of cancer spreading in the other organs in the body.

Diagnosis

Lung cancer symptoms include fatigue, weight loss, cough, dyspnoea, chest pain, haemoptysis, and hoarseness. In the case of lung cancer suspicion, diagnosis tests are performed including chest X-ray and computed tomography which are both used to confirm the suspicion and plan further tissue investigation to guide treatment. Then, the diagnosis and the tumour staging information are assessed by a variety of clinical techniques. Mainly biopsy or needle biopsy are used as the least invasive and safest approaches. Occasionally, when the probability of cancer is very high, curative surgery is performed [8]. When resected tissues are obtained, their morphological features are characterised by histopathological examination described later in section 1.2.3. Finally, the stage of the disease is defined using the TNM (tumour-node-metastasis) classification system, the international staging system for non-small cell carcinoma. The letter T indicates the position of the tumour and its size. For example, T2a represents a tumour situated in the main bronchus with a size between 3 and 5 cm. N represents the nodal stage, determined by the extent of tumour spread to the lymphatic system. N1 refers to hilar nodal metastases and N2 to ipsilateral mediastinal. The diagnosis is crucial to determine the appropriate treatment.

Treatment

The main curative existing treatments are surgery, radiotherapy (including stereotactic ablative radiotherapy) and other less used techniques such as radio-frequency and microwave ablation. The treatment depends on the fitness of the patient as well as the tumour stage and type. However, the wishes of the patient must always be respected. Chemotherapy and radiation therapy are used for advanced stage disease and for small cell carcinoma (mostly) and is generally a palliative (non-curative) treatment. Resection is mostly performed in the early stage of non-small cell carcinoma. The prognosis depends on the tumour type and whether there are any driver mutations, such as those of the epidermal growth factor receptor (EGFR). It also depends on fitness and comorbidity, both captured

by the performance status (PS) of the patient. For good performance status (full active patient), the median survival with chemotherapy is 10 to 12 months.

1.2.2 Skin cancer

The human skin is the largest organ of the human body. Skin is a complex organ with multiple layers, it is composed of three main layers: the epidermis, the dermis, and subcutaneous tissues (Figure 1.2).

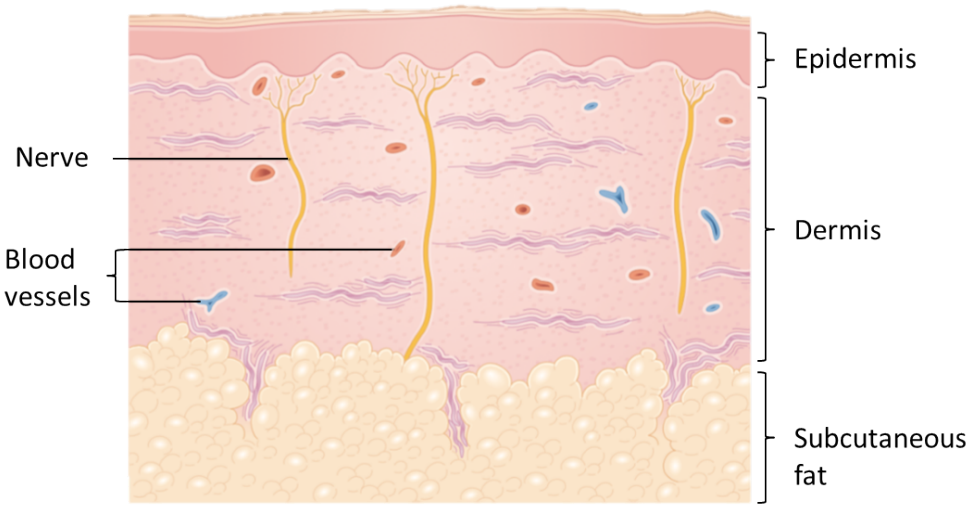


Figure 1.2: Schematic representation of the skin structure (reproduced from [9]).

The epidermis itself is made of layers of epithelial cells. The most superficial layer is the stratum corneum, this layer act as a barrier between the body and the outside world, protecting the body from external substances (chemicals, microbes) penetration but also limiting water loss from underlying skin tissue layers. The epidermis thickness, in average for all body sites, is equal to 84 μm [10]. The dermis is made of two layers of connective tissue that are composed of elastin and collagenous fibres. Its thickness varies with body location, on the face, the dermal thickness lies in the 0.7 to 2 mm range, reaching its minimum value and maximum value respectively for upper medial eyelid and lower nasal sidewall [11]. And finally, the subcutaneous layer, also called the hypodermis, has for function to be an energy reserve in storing the lipids in the fat-storing cells called adipocytes. This normal skin tissue structure described can be altered by

numerous skin conditions but we will only focus our research on non-melanoma skin cancer.

Epidemiology

Skin cancer is divided into two main types: non-melanoma skin cancer, the type with the highest incidence, and melanoma skin cancer, with considerably lower incidence. Melanoma is the 20th most common cause of cancer death in the UK. Melanomas occur mostly in areas exposed to the sun, such as the back and legs, with higher occurrence among light-skinned Caucasians. The risk factors include the predisposition to freckles, family history, skin colour and predominantly intense exposure to ultraviolet (UV) light. Concerning non-melanoma skin cancer, 80 % of the tumours occur on the neck and head. The mortality rate for non-melanoma skin cancers is low, with less than one per cent of total cancer deaths. Death is rare from this type of cancer yet if it happens, it is due to a neglected local invasion or if the cancer has spread to other areas of the body. In the UK, it has had a drastic increase of incidence rate: 150 % since the early 1990s, which is much higher than other common cancers, such as bowel cancer (one per cent), prostate cancer (44 %), breast cancer (20 %). The development of non-melanoma skin cancer is due to a combination of environmental (sun exposure, thinning of the ozone layer), genetic (lack of tanning ability, propensity to freckle), and phenotypic (skin colour) factors.

Pathogenesis

Basal cell carcinoma (BCC) and cutaneous squamous cell carcinoma (SCC) are the two most common subtypes of non-melanoma skin cancer. Basal cell carcinoma is a malignancy of basal cells called keratinocytes found in the basal layer of the skin and responsible for the epidermis formation. It can present different forms from nodular lesions, the most common, to infiltrative, superficial, micronodular, and basosquamous [5]. Squamous cell carcinoma is defined by an atypical proliferation of invasive squamous cells. It occurs on areas of the skin exposed to the sun such as head, hands, legs and forearms. We will exclusively focus on non-melanoma skin cancer later in this thesis.

Diagnosis of non-melanoma skin cancer

Diagnostics tools used in clinical for the screening of non-melanoma skin cancer include dermoscopy, it is a diagnostic technique based on cross-polarized light shining through a magnifying lens directed towards the skin. It helps to visualize colour and microstructure of skin lesions for their discrimination from benign to malignant. For better information on skin morphology, high frequency ultrasonography is used with tuned ultrasound frequencies for different penetration depth and resolution combination. More recently, novel techniques for real-time *in vivo* images of the detailed skin microstructure were developed such as optical coherence tomography (OCT), confocal laser scanning microscopy (FLSM) and reflectance mode confocal microscopy (RCM) [12].

Surgical excision

Most cases worldwide of basal cell carcinoma are treated by surgical excisions. In the case of surgery, if the clinical margins are poorly defined, or if the tumour is bigger than 2 cm, or if a previous treatment on the patient failed, there is a risk for BCC recurrence. For high-risk BCC, Mohs micrographic surgery is the adopted surgery which associate resection and histological diagnosis of the margins with 97 % 5-year cure rates and excellent cosmetic outcomes. This surgery is also called margin controlled excision, it requires special equipment and highly trained surgeons and pathologists. In this thesis, we were only interested in basal cell carcinoma tissue obtained from Mohs surgery which will only be discussed henceforth.

1.2.3 Screening and diagnosis tests

Lung cancer screening tests to detect disease indicator are conducted with chest X-ray or computed tomography while diagnosis tests with histopathology give a definite diagnosis.

Low-dose computed tomography (CT)

This imaging technique uses x-ray as the ionizing radiation; the x-ray beam is continuously rotating in a spiral or helical motion. The patient is placed between the beam and the detector and multiple cross-sectional pictures of the lungs are

taken and processed by the computer. The entire series of pictures can be looked individually as a 2D picture or its in entirety (3D picture). It uses no dyes or injection making it a non-invasive approach for lung cancer screening. For the patient, it is quick and painless. A CT image is made up of pixels of different levels of gray depending on the density of the material or the linear attenuation coefficient of the tissue. Air and soft tissue have the maximal X-ray transmission so they appear the darkest while blood vessels or bone have the maximal X-ray absorption and appear the whitest in the CT image. From the 2D images, the aim is to detect and identify pulmonary or lung nodules. The nodule appears as a white spot on the CT images. The nodule size is an important characteristic to define the risk of malignancy, its shape, and growth rate (difference of size over time) are parameters to differentiate it being benign from malignant [13].

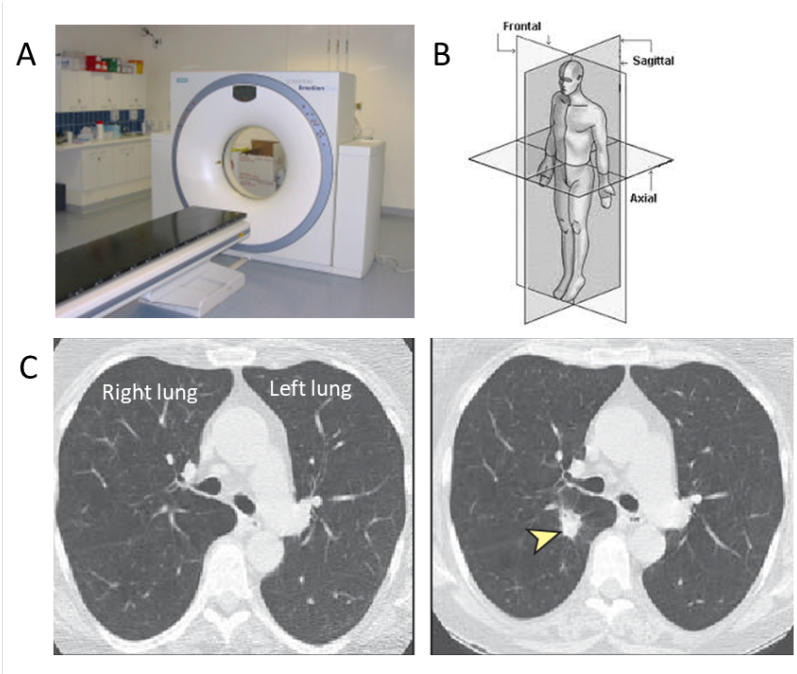


Figure 1.3: Low-dose computed tomography (CT): **A** photo of a CT scanner, **B** reference planes for image capture, **C** low-dose computed tomography scans showing the apparition of a nodules between two scans taken a year apart. The nodule is indicated with a yellow arrow (from[14]).

Chest X-ray radiography

This conventional radiology technique has become less used for lung screening in favour of computed tomography due to the major advances in computer

technology leading to an increased resolution for nodule detection in CT scans. Chest X-ray is based on a similar principle using the same source of ionizing radiation. The image density and shade are similar to CT scans thus tissue structures follow the same behaviour. The X-ray radiograph has a low cost making it easy to access for health care facilities. The major difference with CT is that a chest x-ray provides a 2D image only. Ultimately, once a potentially malignant nodule is detected, it is taken out from the lung by fine needle biopsy and sent for microscopic examination.

In vivo diagnosis techniques *e.g.* reflectance confocal microscopy and optical coherence tomography have been developed to provide additional information not available by a simple visual inspection of the skin.

Reflectance confocal microscopy (RCM)

This non-invasive imaging technique allows investigating intra-cellular structures of the skin with a typical lateral resolution of 5 μm. Laser light is focused onto the skin where reflection occurs when the light interacts with two cellular structures of different indexes of refraction. The reflected light from the focal point travels back towards the objective lens that focuses it into a small pinhole aperture. As a result, the light coming from out-of-focus planes is masked out by the pinhole. By changing the focus in the vertical plane (z-direction), a series of stack images is obtained [15].

Optical coherence tomography (OCT)

In OCT, a 1310 nm light source is used to distinguish skin structures of different refractive indexes. The light passes through a beam splitter to be split into a reference beam and a beam directed at the skin. When the two beams recombine, if the coherence is maintained, they will interfere at the detector. The grayscale images recorded indicate structures with different contrast. Compared to RCM, this technique has the advantage to be able to image up to 2 mm in the tissue however with a lower resolution of approximately 8 μm.

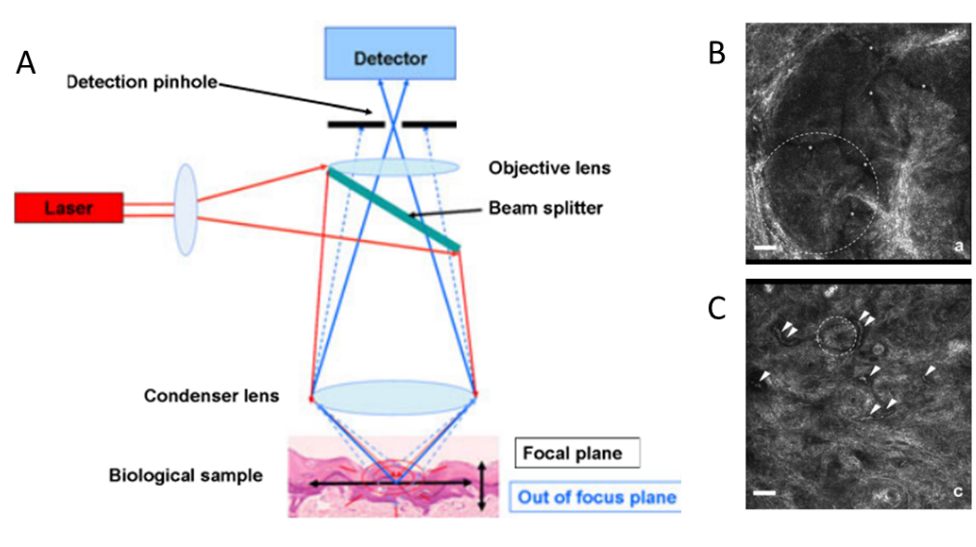


Figure 1.4: **A** Diagram illustrating the confocal scanning laser microscopy technology, Reflectance confocal microscopy image showing **B** islands of BCC tumour cells with high refractivity, and **C** dilated blood vessels (arrowheads) surrounding tumor cells (dotted circle). Reproduced from[15] (scale bar: 50 μm).

Histopathology

The current gold method for cancer diagnosis, for both skin and lung tissues, is histopathology. Histopathology consists in the examination of the stained tissues for the characterisation of their morphological features. Sectioned tissues resulting from biopsy or surgery, for lung and skin respectively, are subjected this microscopic examination. The most widely used stain combination for histopathology is haematoxylin and eosin, refered as H&E. The process of staining is done in a very controlled environment by specifically trained staff. It is because staining time, pH and concentration of the dye solution are critical factors to obtain H&E stained slides meeting the quality requirements for proper disease screening. The image quality and contrast of the H&E stained slides microscopic images allow distinguishing various morphological features used for histopathological diagnosis. The H&E staining is made from haematoxylin which has a deep purple-blue colour and stains nucleic acids, and eosin which has a pink colour and stains proteins. Haematoxylin is a natural dye obtained from the Haematoxylum campechianum tree. Haematoxylin is oxidised by a chemical oxidant to haematein, a molecule that binds to metal ions such as aluminium (Al^{3+}) and iron (Fe^{3+}). The haematein-metal complexes created have a positive charge thus the haematein/aluminium is likely to be attracted to negatively charged groups such

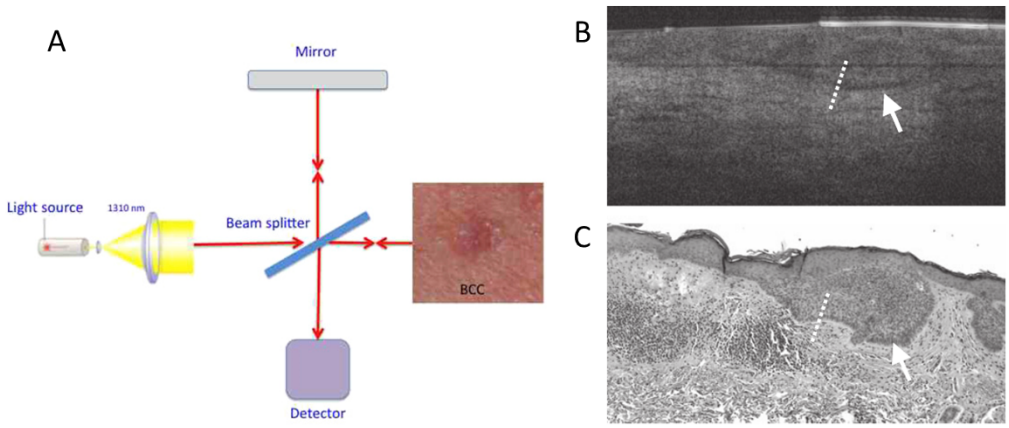


Figure 1.5: **A** Schematic representation of the optical coherence tomography setup. **B** OCT image of a superficial BCC indicated by the darkening of the epidermis region (white arrow), **C** H&E confirming the diagnosis, closely packed cells grouped in a nest (white arrow) and the surrounding stroma. Reproduced from[16] and from [17].

as chromatin in cell nuclei. On the other side, eosin is an orange-pink synthetic dye negatively charged which, by electrostatic interaction, will bind to positively charged proteins in tissue. In general, eosin is staining the rest of the unstained tissue with shades of pink from lighter shade for collagen to brighter shades. Eosin binds to cationic proteins such as arginine and lysine in the cytoplasm and connective tissue creating a bright pink colour. It is also strongly absorbed by cationic nuclear proteins in red blood cells generating a red-orange shade.

1.2.4 Limitations of the standard diagnosis methods

Several limiting factors for the previously described techniques are discussed in this section. They are mainly due to high costs, extended time to perform the procedures, and the training required for the hospital staff. For skin cancer, firstly, the screening by RCM or OCT is challenging for dermatologists due to the difficulty in the interpretation of the grayscale images. The recognition of the morphological features is only possible for highly experienced users. Then, the diagnosis is provided from the tissue resection; itself obtained from Mohs micrographic surgery. Mohs surgery is a complex surgical procedure with the need of specialist surgeons, technicians, and treatment facilities. For lung cancer, similarly, the surgery leading to the diagnosis of tissue biopsy requires appropriate

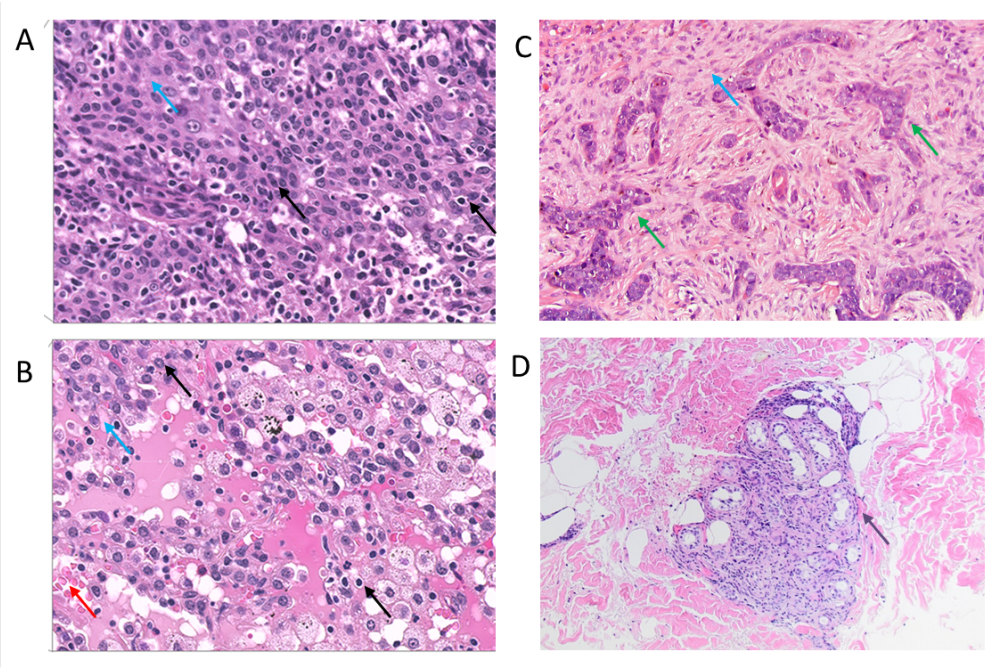


Figure 1.6: Representative stained specimens showing **A** squamous cell carcinoma, and **B** adenocarcinoma for lung cancer, and **C** BCC (green arrow), and **D** dermis with inflammation (purple arrow) for skin cancer (From [18]). The nuclei with their high content in nuclei acid are stained purple (black arrow), while the connective tissue (blue arrow) is pink. The red blood cell, carrier of oxygen, are reddish.

equipment and time availability while it can lead to pain and complications for the patient. Then, it is a time-consuming technical challenge to remove and reinsert the biopsy tool to collect more than one tissue sample. On top of that, for both lung screening technologies, there is a non negligible number of false positive cases. Overdiagnosis is defined as the detection while screening of a cancer that would not have been clinically apparent without the test. The overdiagnosis due to false positive from CT scanning or X-ray may lead to unnecessary secondary test and patient anxiety. Also, once again, the process of detecting disease from the apparent nodules remains dependent upon human observation and interpretation [19].

For both cancer types, many tissue resections can be taken in one session, but the procedure leading to diagnosis of the tissues is time consuming and therefore expensive and is not excluding risk for the patient.

Concerning histopathology, preparing tissue specimens for histological examination requires care and skill. From specimen collection and transport to staining, a dozen strict steps are to be followed for grossing, processing, embedding, and for

microtomy and flotation [20]. Even with a highly trained specialist, H&E staining can lead to errors slowing down the process of diagnosis. As an illustration, we show some examples of the commonly encountered errors in Figure 1.7.

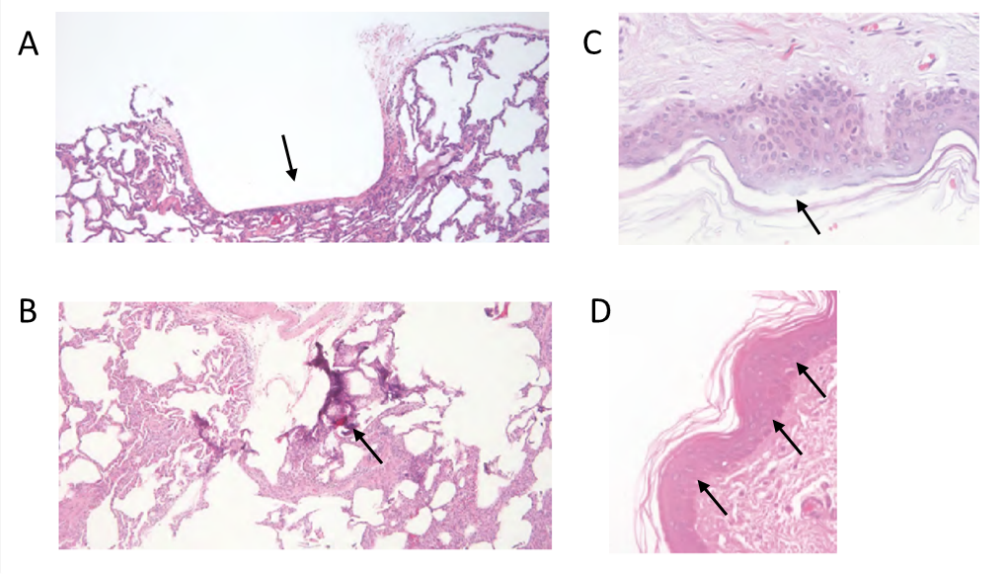


Figure 1.7: Typical examples of poor H&E staining results for lung tissues **A** local trauma from rough handling, **B** section of tissue sticking to another part of adjacent tissue, and for skin tissues, **C** showing uneven haematoxylin staining seen in the epidermis region, and **D** representing over-stained with eosin epidermal nuclei appearing pinkish instead of purple. The errors are highlighted with a black arrow. Reproduced from[20].

Such inappropriate results from H&E staining can lead to an incorrect diagnosis or a delay in the diagnosis if an additional staining procedure is necessary.

Then for microscopic assignment, for lung and skin cancer types, the methods are prone to subjectivity and intra-user variation induced by human error. Indeed, some morphological and structural features assessed can be confounded. As an example, typically for skin, inflamed dermis and epidermis appear both as purple regions in the H&E slides, due to an equivalent nucleus to cytoplasm ratio making them hard to distinguish with human judgment only.

In this context, we present the development of an objective method, based on Raman spectroscopy, to assess molecular differences between tissues that could be implemented in the diagnosis routine.

1.3 Raman spectroscopy (RS)

Theory of inelastic scattering

The phenomenon of inelastic scattering of light known as the Raman effect was discovered by Sir Chandrasekhara Venkata Raman for which he was awarded the Nobel Prize for Physics in 1930. Sir Raman observed a shift in frequency of scattered photons compared to the frequency of the excitation radiation, which he showed was due to the vibrational and rotational proprieties of the molecules in the sample. The advances in technology brought up in 1960, the technique of Light Amplification by Stimulated Emission of Radiation known as the laser, an advanced device capable of producing an intense concentrated beam of monochromatic light [21]. It came as a perfect light source for the advance of the spectroscopy techniques. These techniques, relying on the collection of photons, required improvement in the collection devices in terms of speed of collection, capacity of storage and sensitivity to incoming photons. In 1969, the invention of the charge-coupled device (CCD) was a major technological advance and revolutionised the detection in Raman spectroscopy [22]. This multichannel array detector is extremely sensitive and can capture weak optical signal converting tens of photons into an electronical charge. Soon after the technology was made available, the use of Raman spectroscopy for examining biological systems has been very popular. Indeed, it is a non-destructive method capable of providing information about vibrational and rotational states of biological systems without extrinsic labelling.

The origin of the Raman scattered radiation comes from the interaction between the electromagnetic field from the incident laser radiation and the oscillations of the electric moments of the electrons in molecules. Raman shift corresponds to vibrational energy levels of the molecules of the sample measured and so on the type of atom and type of chemical bonds between atoms. As a convention, the Raman shift $\bar{\nu}$ is usually expressed in terms of wavenumbers in reciprocal centimetres units (cm^{-1}). The photon energy E is given by

$$E = hc\tilde{\nu} = h\nu \tag{1.1}$$

with the photon frequency $\nu = \omega/2\pi$, h the Planck's constant and c the speed of light in vacuum.

When irradiating a sample with a laser wavelength away from any absorption band, it will excite a molecule by $h\nu$ where h is the Planck's constant and ν is the frequency of the laser light. We expect the distribution of energies of a molecule to follow the Maxwell- Boltzmann distribution stating that the ratio of the population of excited states N_f to ground state N_i varies as

$$\frac{N_f}{N_i} = exp\left(-\frac{\Delta \mathrm{E}}{K_B T}\right) \tag{1.2}$$

With k_b the Boltzmann constant (J.K^{-1}), T the temperature (K) and $\Delta E = \varepsilon_f - \varepsilon_i$ the energy difference between the level f and the level i. It means that under normal conditions, there are exponentially fewer molecules starting at a higher energy vibrational state. It implies that most molecules are in the vibrational ground electronic state $S_{0,\nu=0}$ and few are in the next higher vibrational level $S_{0,\nu=1}$. Raman scattering occurs when the scattered photon is of a different energy than the incident photon. The photons can be of eigher higher or lower energy depending on the vibrational state of the molecule under study leading to Stokes or anti-Stokes radiation. Stokes scattering arise when instead of returning to the ground state, the molecule returns to a vibrationally excited state after its excitation. It leads to a change in energy corresponding to the difference between the two states. Stokes scattering occurs when energy is transferred from the incident photon to the molecule. This process induces an increase of the amplitude of the molecular vibrational mode. On the other hand, Anti-Stokes scattering happens when following the excitation, the molecule returns to the ground state with the emission of a photon with higher energy. As demonstrated from equation (1.2), since most molecules are found to be in the ground state at room temperature, Stokes scattering is more frequent that anti-Stokes scattering. Figure 1.8 gives a schematic representation of the considerations discussed above.

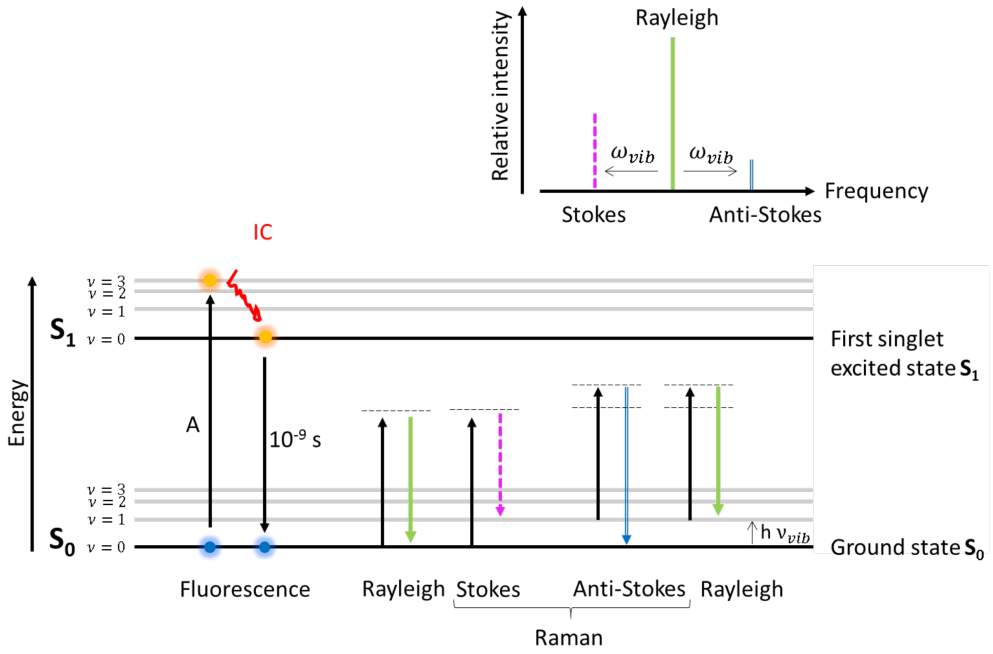


Figure 1.8: Jablonski diagram showing the vibrational transitions of energy involved in the events of fluorescence, Rayleigh scattering and Raman scattering. The thick lines represent the electronic ground state S_0 and the first excited electronic state S_1. The process of fluorescence emission involves photon absorption (A), internal conversion (IC) and relaxation of the photons from singlet state S_1 back down to the ground vibrational level ν_0 of the ground electronic state S_0. During Rayleigh scattering, molecules in the virtual state of energy will relax back down to $S_{0,\nu=0}$ or $S_{0,\nu=1}$ with energy $h\nu_0$. Raman scattering implicate virtual states (dotted lines), for Stokes scattering molecules in the virtual state $h(\nu_0 + \nu_{vib})$ relax to the vibrational level $\nu = 1$ generating scattered photons at a frequency equal to $\nu_0 - \nu_{vib}$. On the other hand, molecules in virtual state $h(\nu_1 + \nu_{vib})$ relax to the ground vibrational state $\nu = 0$ generating scattered photons at a frequency equal to $\nu_0 + \nu_{vib}$. We note that the rotational energy levels are not represented on this diagram.

Using a classical electromagnetism model, the incident monochromatic laser radiation of frequency ω_L on a diatomic molecule has an electric field written as

$$\vec{E} = \vec{E}_0 cos(\omega_L t) \tag{1.3}$$

which create a displacement of the electron cloud of the molecule changing its polarizability α. The interaction between the molecule and the electromagnetic wave generates an induced dipole moment of the bond $\vec{P}_{in}$ such as

$$\vec{P}_{in} = \alpha \cdot \vec{E} = \alpha \cdot \vec{E}_0 cos(\omega_L t) \tag{1.4}$$

The perturbation from the normal vibration of the molecule will bring the molecule to a different vibration state that can be described as a combination of normal modes q such as

$$q^{(m)}(t) = q_0^{(m)} cos\left(2\pi\nu_{vib}^{(m)} t\right) \tag{1.5}$$

with $\nu_{vib} = \frac{\omega_{vib}}{2\pi}$ the harmonic frequency of the m$^{\text{th}}$ vibrational coordinate. Since we consider a diatomic molecule that present only one axis of vibration, m can be ignored. In general, m will represent the possible translations and rotations of every atoms in the molecule meaning $3N-6$ (or $3N-5$ for a linear molecule) vibrational degrees of freedom for a molecule of N atoms. The polarizability α can now be expressed in terms of this displacement as a Taylor series expansion

$$\alpha(q) = \alpha(q_0) + q \cdot \left.\frac{\partial\alpha}{\partial q}\right|_{q=q_0} + \frac{q^2}{2} \cdot \left.\frac{\partial^2\alpha}{\partial q^2}\right|_{q=q_0} + \dots \tag{1.6}$$

By substituting α in equation (1.2), ignoring the higher order terms, taking $\alpha(q_0) = \alpha_0$ and noting the following trigonometric identity $cos(A)\,cos(B) = \frac{1}{2}\left[cos(A+B) + cos(A-B)\right]$, we have

$$\begin{aligned}
\vec{P}_{in} &= \alpha_0 \cdot \vec{E} + q \cdot \vec{E} \left.\frac{\partial\alpha}{\partial q}\right|_{q=q_0} \\
&= \alpha_0 \cdot \vec{E}_0 cos(\omega_L t) + q_0 \cdot \vec{E}_0\, cos(\omega_{vib} t)\, cos(\omega_L t) \left.\frac{\partial\alpha}{\partial q}\right|_{q=q_0} \\
&= \alpha_0 \cdot \vec{E}_0 cos(\omega_L t) + \frac{1}{2} q_0 \vec{E}_0 \left[cos(\left[\omega_L + \omega_{vib}\right] t) + cos(\left[\omega_L - \omega_{vib}\right] t)\right] \left.\frac{\partial\alpha}{\partial q}\right|_{q=q_0}
\end{aligned} \tag{1.7}$$

From equation (1.7), we deduce that, by treating the system classically, the molecule will radiate light at three frequency components. The first component represents the Rayleigh scattering which has the same frequency ω_L as the incident laser radiation and a magnitude proportional to the polarizability of the molecule α_0. The second and third terms are components of the Raman scattering corresponding to anti-Stokes scattering $\omega_L + \omega_{vib}$ and Stokes scattering $\omega_L - \omega_{vib}$. Unlike Rayleigh scattering, for Raman scattering, the photon is scattered at a different energy (frequency and wavelength) than the incident photon corresponding to an energy loss for Stokes and a gain in energy for anti-Stokes.

In this thesis, we will only consider the Stokes-Raman scattering later abbreviated as Raman scattering. Additionally, equation (1.5) highlights the fact that the condition for observation of Raman scattering is a change of the molecule polarizability when it vibrates in the electric field such as

$$\left.\frac{\partial \alpha}{\partial q}\right|_{q=q_0} \neq 0 \tag{1.8}$$

Equation (1.8) is the primary selection rule for Raman scattering.

1.4 Polarised Raman spectroscopy (PRS)

Raman scattering has the particularity to be sensitive to the polarisation of the incident light. If the incident laser light is linearly polarised, it will induce most of the Raman signal to be polarised in the same plane as the incident electrical field. It is because $\vec{P}_{in}$ from equation (1.4) is, in fact, a tensor of rank 2, thus the equation can be rewritten as

$$\vec{P}_{in} = \begin{bmatrix} P_x \\ P_y \\ P_z \end{bmatrix} = \begin{bmatrix} \alpha_{xx} & \alpha_{xy} & \alpha_{xz} \\ \alpha_{yx} & \alpha_{yy} & \alpha_{yz} \\ \alpha_{zx} & \alpha_{zy} & \alpha_{zz} \end{bmatrix} \begin{bmatrix} E_x \\ E_y \\ E_z \end{bmatrix} \tag{1.9}$$

where P_i (i = x, y or z) is the oscillating induced moment in the i$^{\text{th}}$ direction, α_{ij} are the polarizability derivatives and E_i is the electromagnetic field component in the i$^{\text{th}}$ direction. In this notation x, y, and z are the molecule-fixed coordinates. The equation (1.9) illustrates the connection between the polarisation of the scattered radiation and the polarisation of the incident radiation [23].
For the Raman vibrational transition n $\rightarrow$ m denoted as $\sigma_{n,m}$ the Raman cross-section is defined by

$$I_{n,m} = \sigma_{n,m} I_0 \tag{1.10}$$

where I_0 is the intensity of the incident radiation and $I_{n,m}$ is the scattered intensity integrated over all angles and polarisation directions for a non-oriented sample

[24]. The Raman cross section is connected with the Raman polarisability as follow

$$\sigma_{n,m} \propto (\nu_0 \pm \nu_k)^4 \cdot \sum_{\rho,\sigma} |\alpha_{\rho,\sigma}|^2 \tag{1.11}$$

The indices ρ, σ correspond to the molecule-fixed coordinates. The equation (1.11) demonstrate that since $\alpha_{\rho,\sigma}$ depends on the orientation of the molecule, there is a direct correlation between the orientation and the polarisation of the Raman scattered light. We can define two components for the Raman scattered radiation intensity I_{par} and I_{perp} respectively polarised parallel and perpendicular to the electric field vector of the incident light. If we consider a system where the electric field vector is in the z-axis direction, the propagation of the laser light in the y-axis direction and the detection of the scattered light is made in the x-axis direction, the two components are written as

$$\begin{aligned} I_{par} &\propto |\alpha_{z,z}|^2 \cdot \left|\vec{E}_z\right|^2 \\ I_{perp} &\propto |\alpha_{y,z}|^2 \cdot \left|\vec{E}_z\right|^2 \end{aligned} \tag{1.12}$$

The depolarisation ratio ρ defined by the ratio of the perpendicular polarised and the parallel polarised Raman scattered light, we obtain

$$\rho = \frac{I_{par}}{I_{perp}} = \frac{3\gamma_s^2 + 5\gamma_{as}^2}{45\bar{\alpha}^2 + 4\gamma_s^2} \tag{1.13}$$

with

$$\begin{aligned} \bar{\alpha} &= \frac{1}{3}\sum_{\rho\rho} \alpha_{\rho\rho} \\ \gamma_s^2 &= \frac{1}{2}\sum_{\rho\sigma}(\alpha_{\rho\rho} - \alpha_{\sigma\sigma})^2 + \frac{3}{4}\sum_{\rho\sigma}(\alpha_{\rho\sigma} - \alpha_{\sigma\rho})^2 \\ \gamma_{as}^2 &= \frac{3}{4}\sum_{\rho\sigma}(\alpha_{\rho\sigma} - \alpha_{\sigma\rho})^2 \end{aligned} \tag{1.14}$$

where $\bar{\alpha}$ is the mean polarisability and γ_s and γ_{as} the symmetric and antisymmetric anisotropy respectively. For a totally symmetric molecular vibration, the Raman scattered light will retain the incident polarisation leading to a higher intensity detected in the parallel direction. The vibration will lead to a polarised band since it preserves molecular symmetry. From the definition of the depolarisation ratio, we can simplify by saying that a small depolarisation ratio value

indicates a high degree of molecular alignment in the sample. For bands arising from total symmetrical electronic Raman transition $0 \leq \rho < \frac{3}{4}$ and from non-totally symmetric transitions $\rho = \frac{3}{4}$ [25]. From the depolarisation ratio can be derived important chemical information about the symmetry of vibrational modes [26]. The measurement of polarised Raman spectra also can help for the assignment of peaks and reflect different molecule organisation within the sample.

1.5 Literature review

Diseases such as cancer are induced by biochemical changes in tissue. Since the 19th century, tissue microscopic changes have been examined for disease recognition [27]. The main hypothesis of the research presented in this thesis is that the advances in technology (laser, optics, detection technologies) are making biomedical optic spectroscopy the future for diagnosis. Indeed, spectroscopy methods have the capability of providing detailed microscopic tissue information with no damage to the tissue. Fluorescence spectroscopy, elastic scattering, Raman spectroscopy, and Fourier-transform infrared spectroscopy (FTIR) imaging have been developed to analyse the spectrum from the scattered light for clinical diagnosis purposes. These scattering techniques include elastic or inelastic processes. In elastic scattering, the incident photon and scattered photon have the same energy. This type of scattering is taking place for elastic-scattering spectroscopy. In this case, the elastic-scattered light detected represents both scattering and absorption properties of the tissue. The technique has good sensitivity to the absorption bands of tissue components over a wide range of wavelengths. On the contrary, in inelastic scattering, there is an exchange of energy between a photon and the molecule engaged in the scattering. This process takes place in fluorescence spectroscopy imaging for example. This technique is sensitive to the biochemical composition of the tissue. It relies on the detection of the fluorescence signal from the chromophores contained in the tissue. The fluorescent chromophores, also called fluorophores, can be naturally occurring (endogenous) or externally administered (exogeneous). For Fourier-transform infrared spectroscopy, the

different types of chemical bonds present in the molecule are identified by the detection of the infrared absorption spectrum where each absorption or emission peak in the spectrum corresponds to the frequency of a specific functional group. Infrared absorption arises from the direct resonance interaction between the incident radiation and a particular vibration mode of the molecule. This vibrational technique measures absorbance, transmittance, and reflectance of the infrared radiation. In the following sections, we will discuss the development of optical spectroscopy applied to lung and skin tissue characterisation.

Optical spectroscopy for lung cancer imaging and diagnosis

Molecular bioimaging of lung cancer tissues has been achieved in the past using optical spectroscopy methods. *In vivo* methods such as fluorescence spectroscopy imaging are commonly used for diagnosis [28]. Pulsed fluorescence spectroscopy study reports the strong absorption between 380 and 480 nm due to haemoglobin in normal lung tissue when excited at 351 nm [29]. Fluorescence emission of lung tissues excited at 488 nm can discriminate normal tissue from cancerous tissue by the presence of a flavin band at 512 nm and two bands at 554 and 600 nm assigned to flavins and heme proteins [30]. FTIR imaging was performed on normal lung tissues (combined with Raman spectroscopy) [31] and on tumour tissues [32] allowing classification of lung cancer types and subtypes of adenocarcinoma. Besides, malignant and non-malignant lung tissues were discriminated with 96.7 % specificity and sensitivity by attenuated total reflection Fourier transform infrared spectroscopy (ATR FTIR) [33]. The discrimination is based on differences in band positions of protein (1640 cm^{-1}, amide I), carbohydrate (1120 cm^{-1}) and nucleic acid (1085 cm^{-1}) and peak intensities. We can cite several studies that demonstrate the possibility of detecting spectral differences between lung tumour tissues and normal tissue types. Infrared (IR) spectroscopy allows the discrimination between small cell lung carcinoma (SCLC), squamous cell carcinoma (SCC), adenocarcinoma and bronchiolo-alveolar carcinoma which differ in the low wavenumber range (e.g. 400 - 1700 cm^{-1}). Specific deoxyribonucleic acid (DNA) bands can differentiate SCLC from the other cancers. SCC is distinguished from adenocarcinoma by spectral differences associated with glycoproteins in the region with amide I and amide II [34]. In the early

2000s were reported the first studies on near-infrared Raman spectroscopy as a non-invasive diagnosis technique. 1064 nm excitation was used for reduced fluorescence background in the spectrum. The appearance of spectral bands at 1666 cm^{-1} and 1447 cm^{-1} from collagen for tumour tissues was reported as a discrimination between normal tissues and cancerous tissues (immersed in formaldehyde) [35]. The first Raman study at 785 nm of fresh bronchial tissues revealed an extensive number of spectral differences proving the feasibility of this technique [36]. Phenylalanine (1004 cm^{-1}, 1582 cm^{-1} and 1602 cm^{-1}) and tryptophan (752 cm^{-1}, 1208 cm^{-1}, 1552 cm^{-1} and 1618 cm^{-1}) bands have higher intensity, peak from collagen appeared at 1322 cm^{-1} and 1335 cm^{-1} and bands from nucleic acids are wider and larger for malignant tissues. The importance to translate the optical techniques to clinical use has rapidly been an aim for lung diagnosis applications. The lung is a hardly reachable organ constraining its access to flexible and thin fibre optic probes including excitation and collection fibre bundle. A 1.8-mm Raman fibre optic probe was developed [37], showing peak variations for SCC and normal tissue in the high frequency wavenumber range mostly from stretching modes of phospholipids and proteins (2880 cm^{-1} and 2935 cm^{-1}) and spectra dominated by fluorescence in the lower wavenumber range. *Ex vivo* shifted subtracted Raman measurements with an endoscopic probe [38] revealed higher peak intensity from amide I (1655 cm^{-1}), amide III (1260 cm^{-1}) and bending modes of collagen (1300 cm^{-1}) and phenylalanine for tumour tissues.

All these label free techniques allowed the classification of lung cancer using computer algorithms and the aid of histopathology with complete objectivity. They require minimal or no sample preparation. They have their own advantages and disadvantages in terms of spatial resolution, imaging capability, signal strength and cost. Concerning the fluorescence techniques, they only investigate few specific molecular changes in tissues due to a lack of fingerprint specificity. Then, the ATR-FTIR method is limited from the fact that it only distinguishes between malignant and non malignant limiting the relevancy for future clinical diagnosis due to a lack of information necessary to plan a treatment. For FTIR measurements, an extended time was required for the measurement (1 hour to collect 100 spectra) during which tissue might have had a structural modification.

Additionnally, dispertion artefacts were observed in FTIR images that were hard to reduce and correct leading to spectra distortion. In the IR study, spectral differences were too small between tissue types so the authors based their work on second derivative spectra which can increase noise in the signal.
We believe that Raman spectroscopy has the most strength to be adapted as a complementary technique in lung cancer diagnosis. It combines sensitive detection in a non-invasive manner while it has the potential to be miniaturized. Raman spectroscopy has the advantage that absolutely all the molecules exhibit specific spectral signatures, regardless of the excitation laser wavelength, so the molecules in a biological can be investigated all at once. The main weakness of Raman spectroscopy for biological measurements is the slow measurement speed induced by the low laser power limitation (restrained by permissible exposure limit on biological tissue) and the low efficiency of the Raman effect. However, for a use of this method as an add-on to the histopathology routine, an assessment of the tissue region of interest can be done collecting a limited number of spectra respecting a relatively fast diagnosis.

Optical methods for non-melanoma skin cancer research

Diverse techniques are used for skin tissue structure and morphology understanding, fundamental in skin cancer research. First, microscopy techniques like *in vivo* confocal reflectance microscopy showed promising results in diagnosis BCC such as monomorphism of cells, elongated or oval nuclei and increased vascularity [39]. As well as confocal scanning laser microscopy [40]. Although the methods discriminate tissue with high spatial resolution and contrast, they provide information on skin architecture rather than molecular composition. Hence, spectroscopic techniques are preferred for obtaining quantitative information on skin tissue molecular content. Fluorescence spectroscopy is one of them and studies demonstrated its potential for the detection of pathologic tissues. *In vivo* autofluorescence of non-melanoma skin cancer showed an increase emission of tryptophan and an emission decrease of collagen and elastin [41], laser-induced fluorescence spectroscopy at 410 nm excitation was proven to be effective for detection of BCC and SCC [42], multiphoton laser scanning microscopy revealed specific autofluorescence features of SCC and superficial SCC with imaging depth

limitation (up to 200 μm) [43]. However, often in fluorescence studies, the imaging contrast is enhanced with exogenous markers for protoporphyrin IX (PpIX) contrast, such as cream applied to the skin lesion, in bispectral fluorescence measurement [44], or delta-aminolevulinic acid (ALA gel) in digital fluorescence imaging [45]. For non-invasive molecular specificity, vibrational spectroscopy such as IR spectroscopy and Raman spectroscopy are reported in the literature. Raman spectra obtained from skin tissues showed changes in protein (in amide I and amide III regions), lipid (1420 cm^{-1} and 1450 cm^{-1}) and polysaccharide (850 cm^{-1} and 870 cm^{-1}) conformation in BCC compared to healthy skin [46]. And infrared spectroscopy to distinguish malignant from benign tissue was achieved by detection of lipids, carbohydrates and nucleic acids spectral differences [47].

Polarised Raman spectroscopy (PRS) as a tool for characterising biological tissues

For biological studies, PRS is applied for investigation of tissue molecular orientations. It can allow the assessment of preferential molecule vibrations symmetry in tissues [48]. Many studies report the use of such a technique for tissue structure understanding for bone [49, 50], cartilage [51] and collagen orientation in tendon [52–54]. In dentistry, polarised Raman spectroscopy is used as a diagnostic method for early detection of carious enamel by monitoring the polarisation dependence of specific peaks such as 959 cm^{-1} from vibrations of the phosphate ions in hydroxyapatite [55]. In plant science, the dependency of the Raman peak from cellulose peaks at 1095 cm^{-1} determined from PRS allows the understanding of the spatial distribution of cellulose in different plant species. For systems presenting uniaxial symmetry such as silk [56] and DNA fibres [57] molecule conformations and distributions were obtained by polarised Raman measurements. Concerning biological tissues, depolarisation ratio calculated for lipids (2800 cm^{-1} - 2920 cm^{-1}) and proteins (2900 cm^{-1} - 3010 cm^{-1}) and polarised Raman images allowed the determination of specific ordering of breast human tissues revealing modifications in the structure of collagen and lipids for cancerous tissues compared to normal tissues [58]. For cervical cancer, a total of 61 tissue samples was studied showing that a better discrimination sensitivity

and specificity (100 %) was acquired by polarised Raman spectroscopy compared to conventional Raman spectroscopy (respectively 92 % and 75 %). Additionally, the characterisation of changes in DNA orientation (830 cm^{-1}, 1375 cm^{-1} and 1578 cm^{-1}) between normal and cancer cells and collagen degradation (1656 cm^{-1}) in cancer samples was achieved by polarised Raman measurements [59]. Polarised Raman spectroscopy was also shown to provide better discrimination than Raman spectroscopy for basal cell carcinoma and healthy epidermis. The same study displayed depolarisation ratio for phenylalanine (623 cm^{-1} and 1002 cm^{-1}), tyrosine (644 cm^{-1}), polysaccharides (1128 cm^{-1}) and proteins (1450 cm^{-1}) highlighting modifications happening in healthy and peritumoral stroma [60]. Besides, polarised Raman spectra were found to allow better molecular understanding than conventional Raman such as enhanced information on change of collagen conformation (1247 cm^{-1}) and a higher content of DNA in dermis at the proximity of tumour [61].
Polarised Raman spectroscopy has undeniable advantages to be the label-free method for molecule orientation and organisation characterisation for biological tissue studies.

1.6 Aims

The main objective of this work was therefore to establish a lung tissue classification method based on Raman mapping spectroscopy and to prove that its implementation in routine cancer diagnosis can be made possible.

This requires to succeed in the following objectives:

1. Defining from the literature the design of a Raman spectrometer made from optical elements, devices, and opto-mechanical components available off the shelf,
2. Optimising the design taking into account the restrictions applying in hospital premises,

3. Justifying the performance of the system with various standard samples for axis calibration and alignment,
4. Getting familiar with taking measurements from other widely accessible tissues such as human skin tissues.
5. Transporting the system to the hospital and implementing it on the bench,
6. Testing the stability of the system,
7. Working to establish strong connections with histopathologists and staff members from the histopathology laboratory,
8. Understanding the different steps of the H&E diagnosis routine,
9. Supervising the follow up of the collected lung samples,
10. Setting all the parameters for the data acquisition,
11. Defining a consistent and reliable procedure for the spectral measurement of the lung tissues,
12. Optimising data acquisition, analysis, and interpretation.
13. Adapting the design with additional experimental features defined from the results,
14. Establishing a conclusion for the feasibility of lung tissue types differentiation based on molecular differences detected in Raman signal.

2

Development of Raman spectrometer instruments for RS and PRS

In this chapter, we will describe in detail the Raman microspectrometer and the polarised Raman spectrometer setups designed for the study. It includes the steps involved in the calibration of both spectrometers, which are tested for several samples attesting the maximal optimisation. The chapter presents the main components of the instruments, including hardware and software.

2.1 Description of the Raman micro spectrometer

The spectrometer described in this section was designed to be integrated to a room close to the histopathology department at Nottingham City Hospital. We wanted the shortest distance for the lung tissue to travel from the moment it is taken out of the patient's body, the moment we measure it and the moment it is fixed by technicians for its later histopathology assesment. With this in mind we had the opportunity to get a space in the room adjacent to the laboratory in a multi-user environment. Restrictions in terms of space, safety, access, and schedule had to be taken in consideration.

The assigned laboratory bench could not be modified, so no optical table could be used to reduce ambient vibrations. We had the option, however, to add on top of the existing bench a breadboard with mounting holes allowing us to build and robustly fix the optical setup to it. The spectrometer was first assembled at

university where we performed calibration tests and optimisation to make sure we could reach a compact enough design to fit on the hospital table and the chosen breadboard. Besides there was no restricted access to the room where we performed the experiments meaning that to protect people accessing the room while a measurement was ongoing, we designed a full system enclosure built from enclosure panels and laser safety fabric. Also, we kept the laser key and switching-on procedure in a safe with key access.

2.1.1 Instrumentation

The Raman spectrometer was designed based on instruments developed in the group in the past. Components such as laser, spectrograph, camera, and microscope were either already available from previous projects or ordered for this project. Regarding all other components from optical components (lens, mirror, filter) and optomechanical components (cage system, rods, lens mounts) they were all specifically selected for our application.
The optical instrument consists of a monochromatic laser passing through a combination of lenses, mirrors, reaching a microscope objective focusing the laser beam on a sample placed on a transparent substrate. The scattered light is filtered and collected by an optical fibre that directs the radiation to the spectrometer. The Raman spectrometer setup is divided into 4 sections: part 1 represents the laser; part 2 includes the optics in the incident path and the sample, part 3 represent the collection path and part 4 portrays the spectrograph. The schematic representation of the setup is shown in Figure 2.1.

Part 1: laser

In Raman studies, the selection of the laser source is critical; the laser proprieties such as wavelength, mode, stability and power need to be carefully selected. The most widely used laser wavelength (λ) for Raman spectroscopy in biological studies are 532 nm, 785 nm and 1064 nm. The laser wavelength alters the penetration depth of the light in the tissue, the amount of fluorescence generated and has an impact on the excitation efficiency. We justify our choice of wavelength

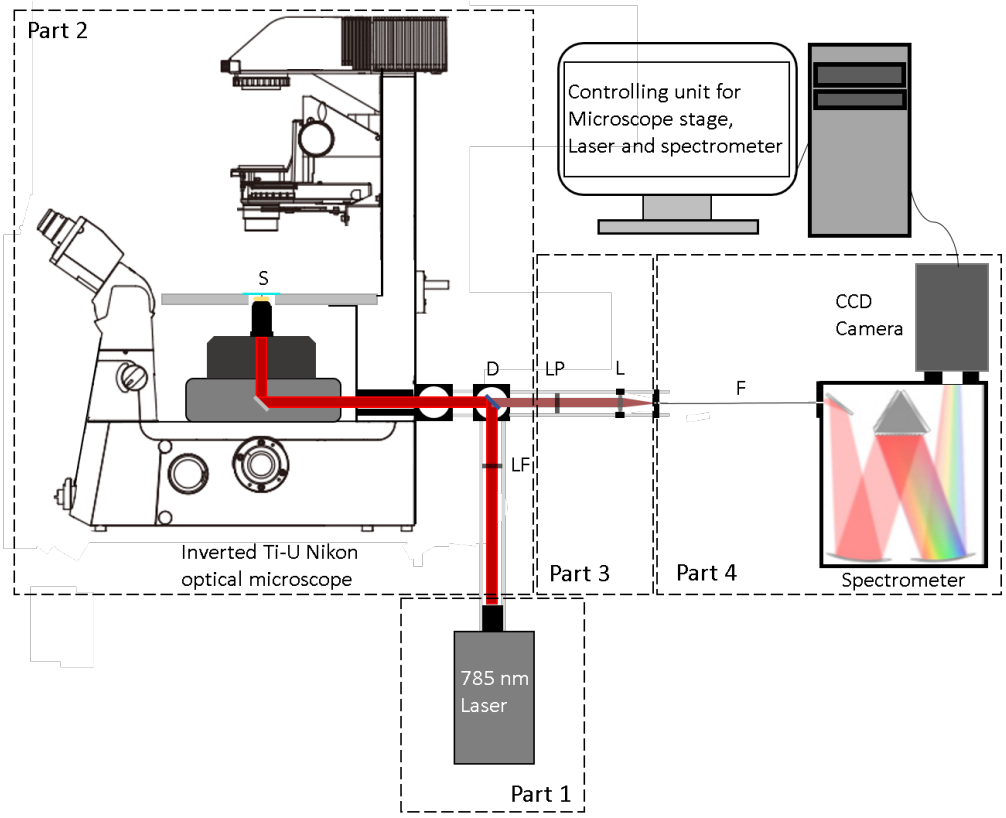


Figure 2.1: Schematic diagram of the Raman spectrometer setup design divided in 4 parts including 785 nm laser clean-up filter (LF), 785 nm dichroic beamsplitter (D), sample holder (S), long pass filter (LP), focusing lens (L) and fibre optic (F).

thanks to the following considerations.

The Raman scattering efficiency is proportional to λ^{-4} meaning a higher efficiency at 532 nm than at 785 nm and at 1064 nm. From this first consideration only it seems that 532 nm would be the preferred wavelength. However, at this shorter wavelength and for organic samples, there is a higher chance to excite electronic transition leading to an increase of the tissue autofluorescence emission. The sample fluorescence creates a background in the Raman spectrum masking the Raman bands and thus better when minimized. For this reason, 785 nm and 1064 nm are preferred over 532 nm. Then, at 785 nm, the Raman scattered signal is higher than at 1064 nm because of a higher Raman scattering cross-section of all molecules. Besides, the chosen wavelength is also required to be in concordance with the CCD sensor responsivity. Most silicon CCDs will be responsive in the

range at which the Raman emission will be for 785 nm. However, for the use of a higher excitation wavelength such as 1064 nm, the silicon CCDs will not be responsive above 1000 nm meaning that InGaAs array detectors will be required, which have significantly higher noise levels compared to Si CCDs.
From every factor taken into consideration, we conclude that 785 nm is the ideal excitation wavelength for our study. The 785 nm laser chosen (XTRA, Toptica) satisfies the criteria of stability, power and performance for our study. It is controlled by an RS232 interface with a threshold current of 900 mA, it provides a 785 nm laser beam of diameter 2.2 mm, working between 15 to 35 °C. Its frequency stability test proved less than 0.1 GHz frequency drift. The laser power at the sample (70 - 160 ± 5 mW) is chosen to be compatible with an application to biological tissues and safe for use in the research laboratory.

Part 2: incident path and sample

The entire system described in this section is built using an optical cage system (Thorlabs) ensuring robustness and accurate alignment. The laser beam radiation is filtered by a laser line filter (LL01-785-12, Semrock) to ensure that the incident radiation is monochromatic, allowing 90 % of the laser light to pass at 785 ± 3 nm for an angle of incidence of 0 ± 2 degrees. The light passes through a dichroic mirror (LPD01-785RU-25, Semrock now LPD02-785RU-25) which specifically reflects the 785 nm incident laser beam at 45 degrees. Then, the beam is directed towards the sample by a mirror placed in the microscope filter cube (fully reflective enhanced silver mirror, Chroma in 91020 filter cube, Nikon) with 97.9 % of transmission at 785 nm at any incidence angle. The light is focused onto the sample through a 60x / 0.85 NA objective oil immersion (RiverD). The numerical aperture (NA) of an objective determines the light-gathering ability via its acceptance cone. The numerical aperture is defined as

$$NA = nsin\theta \tag{2.1}$$

where θ is the half angle of the maximum light cone entering or exiting the objective and n is the refractive index of the medium between the objective lens and the specimen. An objective with a large numerical aperture will have a large

solid angle of collection, so with the use of this kind of objective, we will be able to collect more light consequently providing higher quality Raman spectra. The NA defines the collection efficiency as well as the resolving power and power density of the excitation laser spot. Increasing the Raman signal is possible by increasing laser power (increasing the number of photons) or decreasing the focal volume by using a high NA objective. We have selected an objective with a high numerical aperture with a good transmission, higher than 85 % at our working wavelength, and optimised for fused silica and calcium fluoride substrate 0.1 to 1 mm thick, ideal substrates for tissue sample handling. In addition, the combination of its relatively high numerical aperture and the oil medium implies a small laser spot size for best depth resolution and signal intensity. For each measurement, oil (Cargille) is placed between the objective and the tissue sample, its matching index at 785 nm with the quartz creates a homogeneous path for the laser light to allow full use of the NA. The tissue sample is placed on a quartz coverglass adapted to near-infrared (NIR) Raman spectroscopy with low fluorescence in the fingerprint region compared to glass. The sample sits directly above the objectives on the inverted optical microscope's (Nikon Eclipse Ti-U) motorized XY stage (ProscanIII, Nikon).

Part 3: collection path and camera

The scattered radiation travels through the same dichroic mirror earlier mentioned which lets the longer than 785 nm wavelengths through. It reaches a long-pass filter (LP02-785RU-25, Semrock) used to separate weak Raman signals close to the laser line while blocking the Rayleigh scattered radiations. 93 % of light is transmitted between 795.2 nm and 1770.7 nm. The radiation is focused onto the entrance of the 50-micron diameter optical fibre by a 25 mm focal length lens (AC127-025-B-ML, Thorlabs).

We mount to the side port of the microscope a camera (Lumenera Infinity 2-1) which combines 1.4 megapixel CCD image sensor with 15 fps at full 1392 x 1040 resolution for the collection of low noise high-quality images. It is used to obtain bright-field images of the biological tissue sample and images of the stained sections with the 10x objective (Leica). Both images are used as a guideline for selecting the area in the sample for Raman mapping measurement. The images

captured from the H&E stained tissue sections allowed the identification of a single cell nucleus and provided contrast between the white background and pink shades of the stained tissues. The histopathologist confirmed from the images collected, the image quality required to pronounce a medical diagnosis.

Part 4: spectrometer

The spectrograph is fully automated allowing us to select the most suitable parameters via the user interface. The light scattered from the sample enters the fibre-coupled spectrometer (Oriel 77 200, Newport) through an entrance slit. It then hits the collimating mirror and is directed onto the diffracting grating. Our instrument has a triple grating turret system from which we selected a grating system density of 600 lines/mm over the range 803.2 to 923.6 nm corresponding to wavenumber 288.7 to 1911.7 cm^{-1}. Raman studies of the molecule's vibrations in the high wavenumber (2500 cm^{-1} to 4000 cm^{-1}) region are usually carried to avoid the collection of the sample fluorescent background and from the substrate. However, there is less biochemical information in this wavenumber range than in the fingerprint region. That is the reason why we select the range specified earlier. The light is then reflected from a focusing mirror towards the array detector (Andor iDus DU401 BR-DD). The back-illuminated CCD is optimum for Raman spectroscopy at 785 nm and 830 nm. It can be thermo-electrically cooled to -100 °C to limit dark current (0.0005 e^- /pixel/sec) it combines low-noise electronics for maximized quantum efficiency (up to 95 %).

2.1.2 Calibration

The confocal nature of the Raman microspectrometer designed implies that the laser focal spot is focused onto the sample and the collected scattered light is maximized at the entrance of the optical fibre coupled to the spectrometer.

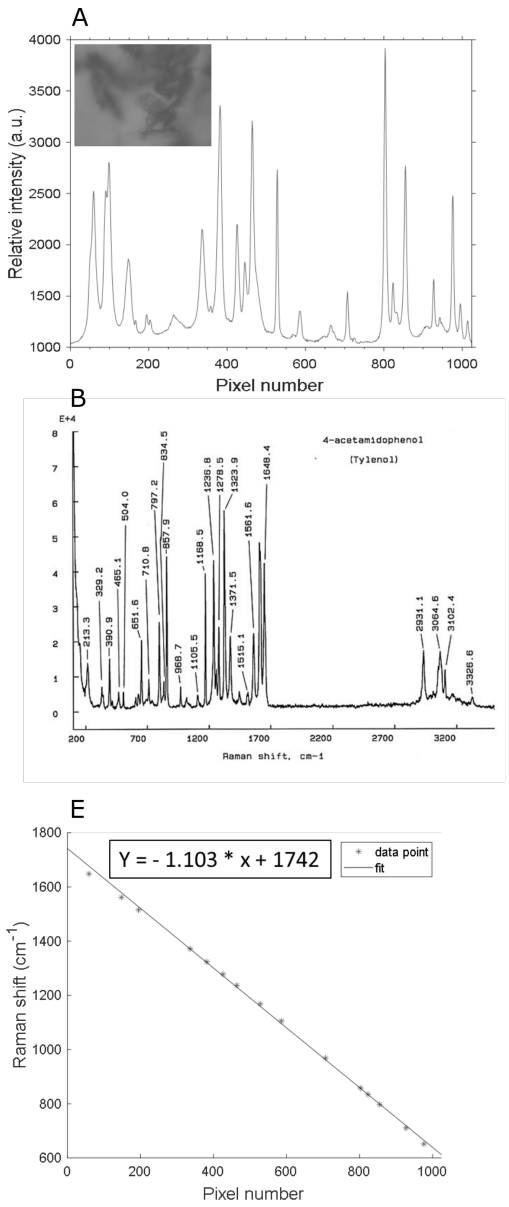
Axis calibration

The American Society for Testing and Materials (ASTM) committee has established Raman shift frequency standards to provide a tool for calibration of the Raman shift axis of the Raman spectrometers [62]. The Raman shift standards were

measured for eight common materials (naphthalene, 1,4 bis (2-methylstyryl) benzene (BMB), sulfur, 50/50 (v/v) toluene/acetonitrile, 4-acetamidophenol (active ingredients of Tylenol), benzonitrile, cyclohexane, and polystyrene) with dispersive Raman spectrometers and Fourier Transform Raman spectrometers from at least six different labs (from the following list FT-Raman, Bruker, FT-Raman, Perkin-Elmer, FT-Raman, Bio-rad, FT-Raman, Dupont, FT-Raman, Nicolet, FT-Raman, Hercules, Multichannel, EIC Laboratories and Scanning (Spex 1403), Ohio State). Each material came from different sources and each lab used the calibration procedure they had in place. For each uncorrected spectrum and each lab, peak Raman shift frequencies were requested for several peaks. The labs had no prior knowledge to the peak frequencies of these materials. Only peak frequencies with standard deviations less than 1.0 cm^{-1} are reported. The values indicated are meant to be used to correct Raman spectra for instrument response functions.

To calibrate our Raman spectrometer, we chose 1,4 bis (2-methylstyryl) benzene (BMB) and 5.4-acetamidophenol (active ingredients of Tylenol) as standards. We first prepared a sample of Tylenol crystals (Sigma Aldrich). Tylenol powder was taken from its plastic container with a clean metal spatula and deposited on the quartz substrate; this operation was done in the chemistry lab under a protected environment to avoid sources of contamination. Then, we focused the laser spot on one of the Tylenol crystal and recorded the Raman scattered spectra. Since the CCD camera has 1024 x 127 active pixels, it means that it records spectra over 1024 columns; we aimed to assign a wavenumber for each column. Because the standard frequencies do not depend on the laser wavelength, the x-axis calibration is performed by comparing the experimental peaks with the standard ones. Fifteen peaks between pixel 0 and peak 1024 were used corresponding to typical standard mean frequencies. The results from Raman measurement on Tylenol are presented in Figure 2.2.

The calibration was fine-tuned with a second step calibration. The method was reproduced for the second sample, BMB, and the final x-axis calibration was established. The BMB sample was prepared similarly that the Tylenol one with the addition of protective equipment (gloves and eye shield). The difference



C

Mean Frequency	Standard Deviation
651.6	0.50
710.8	0.68
797.2	0.48
834.5	0.46
857.9	0.50
968.7	0.60
1105.5	0.27
1168.5	0.65
1236.8	0.46
1278.5	0.45
1323.9	0.46
1371.5	0.11
1515.1	0.70
1561.5	0.52
1648.4	0.50

D

Pixel number	Theoretical Raman shift (cm-1)
59	1648.4
148	1561.6
195	1515.1
336	1371.5
382	1323.8
426	1278.5
464	1236.8
528	1168.5
586	1105.5
707	968.7
803	857.9
823	834.5
855	797.2
927	710.8
976	651.6

Figure 2.2: **A** Spectra of 4-Acetamidophenol with peak wavenumber assignation in **B** and corresponding values of standard deviation in **C** where determined by average of FT-Raman and Scanning data sets from http://www.chem.ualberta.ca/ mccreery/ramanmaterials. Table in **D** correlates the pixel number and its standard Raman shift value. A linear model is fitted to the experimental data points in graph in **E** with the line of best fit equation.

between the theoretical Raman shift frequencies and the standard ones, obtained from the ASTM committee, was used to quantify the overall Raman shift error of our system. The values are reported in Figure 2.3.

A

Tylenol		
Theoretical Raman shift S_{theor} (cm^{-1})	Experimental Raman shift S_{exp} (cm^{-1})	$\|\Delta\|=(S_{exp}-S_{theor})$ (cm^{-1})
651.6	653	1.4
797.2	797.5	0.3
857.9	858.6	0.7
968.7	969.3	0.6
1105.5	1104	1.5
1168.5	1168	0.5
1236.8	1236	0.8
1278.5	1278	0.5
1323.9	1324	0.1
1371.5	1373	1.5
1561.5	1563	1.5
1648.4	1649	0.6
$Sum(\|\Delta_{Tylenol}\|) = 10$		
$n_{\Delta Tylenol} = 12$		
$Mean(\Delta_{Tylenol})=0.83cm^{-1}$		

B

BMB		
Theoretical Raman shift T_{theor} (cm^{-1})	Experimental Raman shift T_{exp} (cm^{-1})	$\|\Delta\|=(T_{exp}-T_{theor})$ (cm^{-1})
642.4	643.1	0.7
950.1	949.9	0.2
978.0	977.3	0.7
1104.1	1104	0.1
1177.7	1178	0.3
1290.7	1290	0.7
1316.9	1316	0.9
1334.5	1334	0.5
1555.2	1555	0.2
1593.1	1594	0.9
1627.9	1628	0.1
$Sum(\|\Delta_{BMB}\|) = 5.3$		
$n_{\Delta BMB} = 1$		
$Mean(\Delta_{BMB})=0.48cm^{-1}$		

C

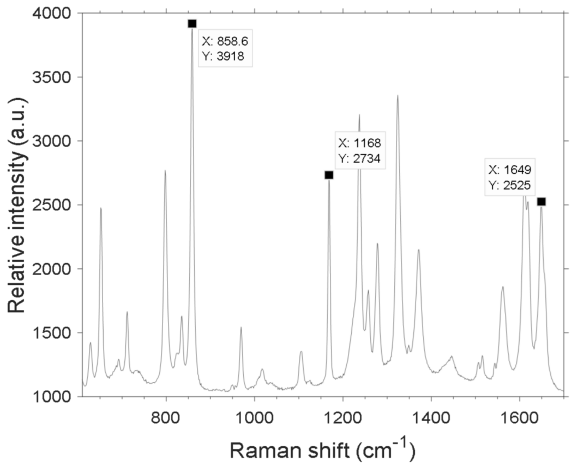


Figure 2.3: Raman spectra in the fingerprint region are recorded from Tylenol and BMB samples. The tables express the variations between the experimental Raman shift values compared to the theoretical values for Tylenol in **A** and BMB in **B**. The spectrum of Tylenol obtained for 2 s integration time is displayed in **C** with three peak values highlighted, illustrating the calibration result.

The average Raman shift error of our Raman hand-built system, calculated from the two calibration samples, is equal to 0.66 cm^{-1}. This value is close to the spectrometer spectral resolution validating the high spectral accuracy obtained. Once, the x-axis had been calibrated, we needed to adjust the automation of the acquisition software. Indeed, the change of the objective between the capture of the microscopic image of the tissue with 4x (or 10x) objective and the Raman mapping measurement (60x oil immersion) require implementing an automated

program that ensures a perfect match of coordinates. To do so, we deposited albumin crystals (Sigma Aldrich) on magnesium fluoride substrate. We selected a region with a crystal with a distinguishable shape. The laser position, represented by a red dot, was chosen to coincide with the centre of the crystal. A bright-field image was captured with the 4x objective. When changing the objective from the 4x to the 60x oil immersion, a manual adjustment of the stage position was needed to realign the laser position and the centre of the crystal. The position drifts in X and Y directions (δX and δY respectively), when the objective magnification is changed, are due to minor parcentricity issues of the microscope meaning that a change of magnification requires refocusing and re-centering. An adjustment of + 122 units in the X-direction (δX = + 122 units) and + 225 units in the Y-direction (δY = + 225 units) was required to perform the alignment. We implemented in the software the values of δX and δY. Then, we selected an area of 1 mm x 1mm with the 4x objective, use the newly developed automatized program and recorded a Raman map. The results are summarized in Figure 2.4.

We observed a perfect correlation between the bright-field image, the principal component analysis image and the 20 clusters Raman map. The spectrum from the two positions mentioned, confirmed the presence of albumin in the crystal and the substrate background in the darker grey region. The spectrum obtained for 1 s integration time clearly shows Raman peaks of albumin highlighted in (f) such as the tyrosine doublet at 828 cm^{-1} and 850 cm^{-1}, the 1003 cm^{-1} phenylalanine band, the 1003 cm^{-1}, the 1447 cm^{-1} CH_2 deformation band and the 1655 cm^{-1} amide I band [63]. While the bottom spectrum showed a peak at 1251 cm^{-1} typical from magnesium fluoride substrate.

Depth resolution

The depth resolution is an important parameter in Raman studies, it indicates how does the optical system perform in effectively collecting the signal at the focal point while rejecting the contribution from the surrounding volume. It depends on instrumental parameters such as the optics and the laser alignment. We have implemented two methods for determining the depth resolution described in the literature [64]. The first method consisted of measuring the

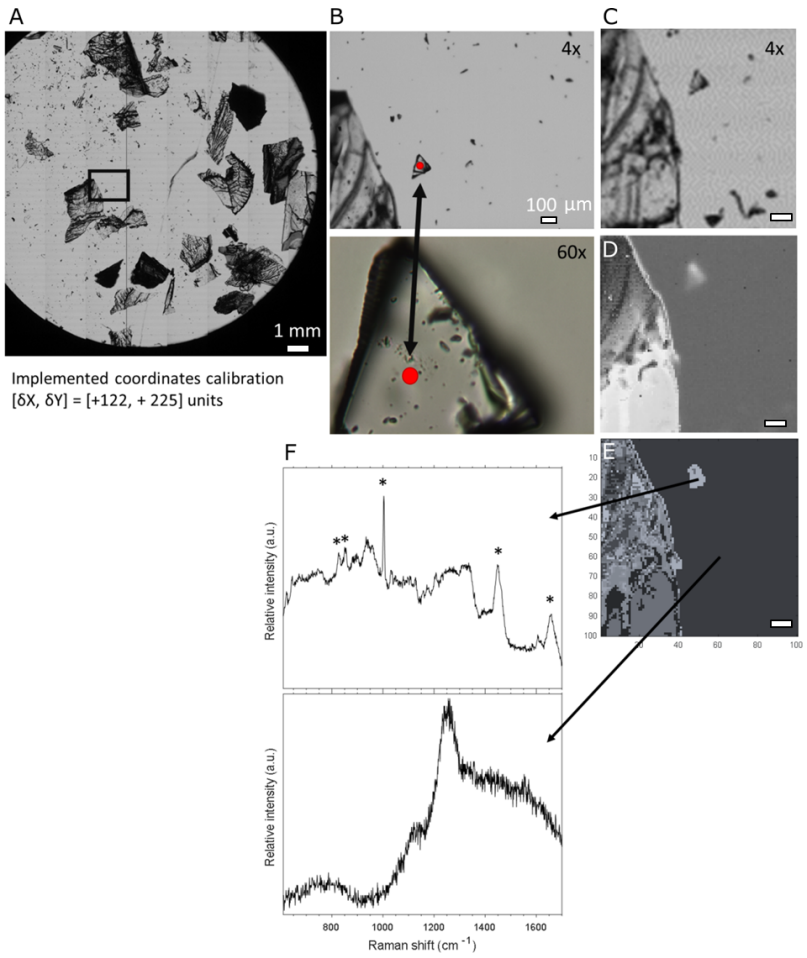
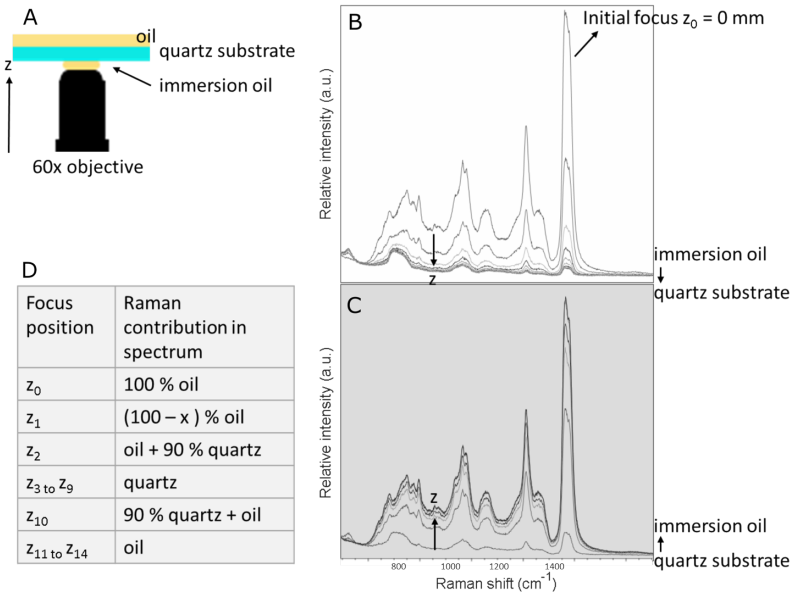


Figure 2.4: A microscopic region from the full sample image in **A** is selected and displayed in **B**. The laser spot is focused at the centre of the triangular crystal of size < 0.1 mm, a manual alignment is performed to obtain the same focus with the 60x objective. The region chosen for the Raman scan is selected **B**, the corresponding PCA image of the Raman map is shown in **D** and a 20-clusters image is obtained in **E**. The spectrum from the regions indicated by the arrows are displayed in **F**.

intensity profile of the Raman signal collected from oil on top of a quartz coverslip. As the schematic of the experiment shows in Figure 2.5, we were interested in studying the evolution of the Raman signal while varying the focus. The initial defocused position starts when we only collected signal from the immersion oil, we then focused higher up into the quartz substrate to get a mixed oil and quartz signal. The signal from the coverslip, distinguished by typical quartz peaks and the decrease of the oil spectral contribution, indicated that the thickness of the coverslip was 0.7 ± 0.1 mm (detected from z_3 to z_9 ± 0.1 mm). This thickness is equivalent to the correct value of the commercial quartz slides meaning that the

optical system successfully collects the Raman signal excited by the laser at the focal point.



D

Focus position	Raman contribution in spectrum
z_0	100 % oil
z_1	(100 – x) % oil
z_2	oil + 90 % quartz
$z_{3\ to}\ z_9$	quartz
z_{10}	90 % quartz + oil
$z_{11\ to}\ z_{14}$	oil

Figure 2.5: **A** Schematic representation of the system used for depth profile measurement with the variable focus direction of the objective possible in the z-direction. First part of the experiment consists in changing by +0.1 mm the laser focus from the immersion oil to the quartz substrate. **B** shows the Raman spectrum (integration time 3 s) for each 10 positions from $z_0 = 0$ mm to $z_9 = 0.9$ mm. The position z_{10} to z_{14} are displayed in **C** while the observations are reported in **D**.

The second method consisted in measuring the full width at half maxima (FWHM) of the depth profile resolution. To do so, we used an identical setup as above and record the spectrum at 3 s integration time every 100 µm. The intensity of the C-H vibrations at 1449 cm^{-1} from oil was used to characterise the depth resolution of our system. The experiment is illustrated in Figure 2.6.

A normal distribution was fitted to the intensity distribution of the band. The standard deviation was calculated to be equal to $\sigma = 67 \pm 2$ µm giving a full width at half maxima equal to FWHM $= 2\,\sigma = 134 \pm 4$ µm. The value of FWHM represents the depth penetration distribution of the laser beam in the sample, FWHM/2 = 67 µm (± 2 µm) is the ideal depth at which the Raman measurement should be collected. This focus requirement is possible and reproducible thanks to the precise wheel focusing mechanism of the microscope that allows a fine focusing of 0.1 mm/rotation with a minimum fine reading of 1 µm.

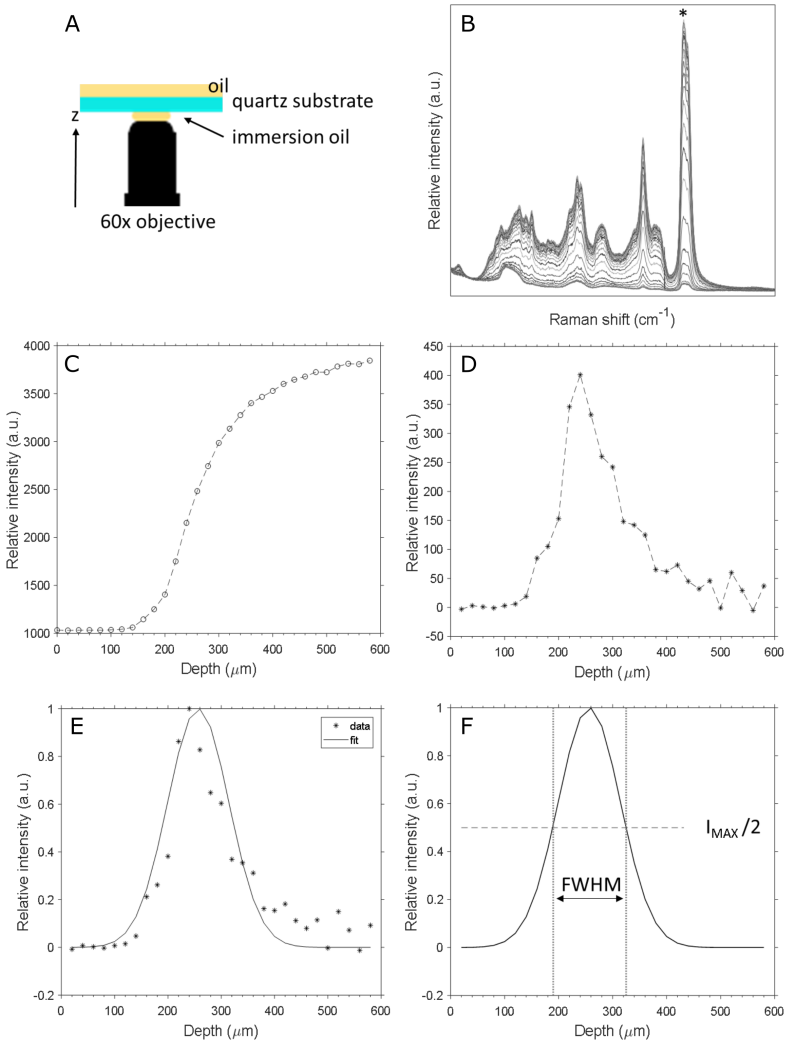


Figure 2.6: **A** Schematic representation of the system used for depth profile measurement with the variable focus direction of the objective possible in the Z-direction. **B** Spectrum recorded at every position with the peak at 1449 cm^{-1} marked by a asterisk. **C** Intensity of the peak at 1449 cm^{-1} as a function of the focus depth. We display in **D** the first derivative of the slope from **C**. A Gaussian function is fitted to the data and displayed in **E** from which we determine the FWHM as shown in **F**.

It should be noted that we are working on resected tissue samples with irregular thickness. In the case of Mohs surgery, the skin tissues are transversely cut creating a smaller thickness on the outer edges of the tissue. For this reason, it was essential to make sure the laser depth penetration agrees with the depth of the tissue within the entire sample. As a result, to ensure consistency between the measurements, the Raman spectra reported in this thesis, are exclusively collected at the ideal depth calculated earlier.

The previous conclusions on depth resolution were tested using 5 μm polystyrene spheres (Sigma Aldrich). We deposited the spheres onto the quartz substrate as represented on the microscopic image in Figure 2.7. We selected a region with two distinct polystyrene spheres from which we recorded a Raman map with a scanning step size of 1 μm for three different focus heights ranging from the surface of the coverslip to 120 μm above it. The signal from polystyrene is expected to be maximal in the region containing the spheres and null everywhere else. From the Raman maps recorded, we generated a 5 clusters maps from which we select the indexes containing spectral vibrations such as peaks at 1001.4 cm^{-1}, 1031.8 cm^{-1}, 1450 cm^{-1}, 1583 cm^{-1} and 1602.3 cm^{-1} exclusively attributed to polystyrene. The measurement showed that the beads are perceived differently (size and position) depending on the focus and so the depth of penetration of our laser light. At the position +80 μm, the polystyrene Raman signal was strong where the spheres are positioned and the clusters showed well-resolved spheres of 5 μm size providing a satisfactory representation of the sphere's shape. However, the defocused position +40 μm and +120 μm showed a distortion of the shape of the spheres. They appeared respectively smaller and larger than their actual size. We concluded from this last depth profile experiment that the position at 80 μm ensures the correct characterisation of the sphere meaning that it is the ideal focus distance in agreement with the laser depth of penetration.

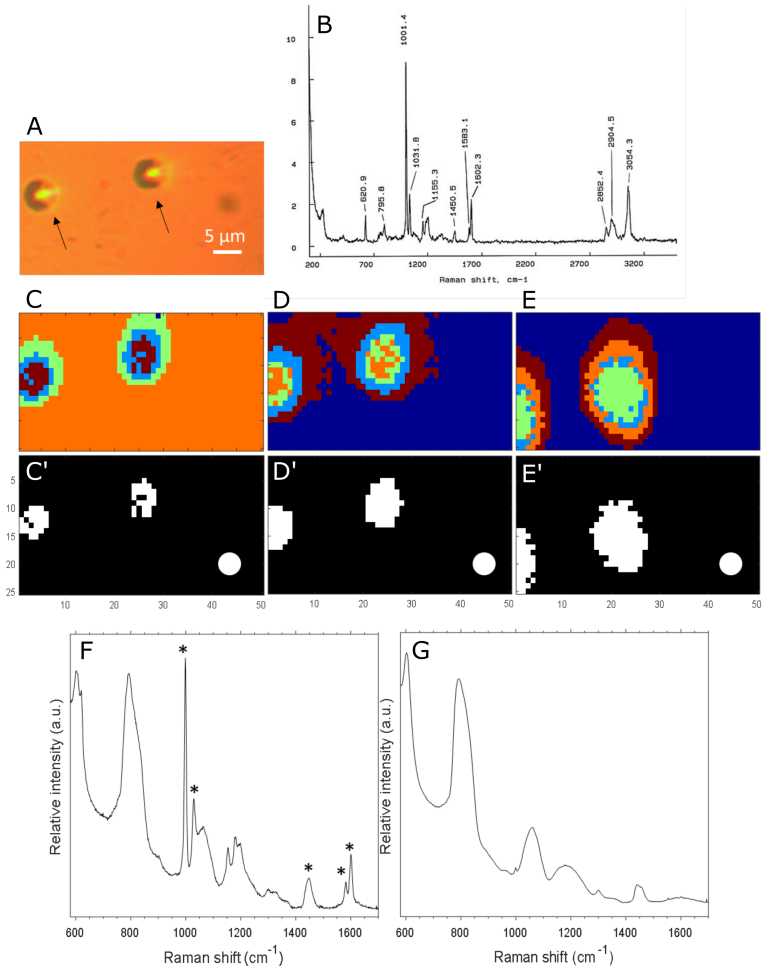


Figure 2.7: **A** Microscopic image of region of interest showing two distinct polystyrene spheres. **B** Theoretical Raman bands from polystyrene. 5 clusters Raman maps of the region for different depth **C** 40 µm, **D** 80 µm, and **E** 120 µm. These clusters are used to generate the black and white colour maps displayed in **C'**, **D'** and **E'**. The white filled circle at the bottom right of each image represent one sphere in real size. **F** represents the Raman spectra in the white region of the 2 clusters image and **G** the spectra in the region represented in black.

2.2 Description of the Polarised Raman spectrometer

As a second part of the project, I worked on the design of a second instrument: a polarised Raman spectrometer. The previously described instrument remained in the hospital premises for other research studies to be conducted there. For the polarised Raman instrument, once again, every component was selected individually and ordered specifically for it. This instrument was installed in one of our research labs within the university. The lab was shared with another PhD student. Similarly to the previous instrument, many aspects had to be taken into consideration in designing the optical bench in this multi-user environment. Excellent stability was obtained thanks to the large optical table we had available in the lab. The laser was shared alternately between the two users. A reflective mirror in a kinematic magnetic mount plate was used by the other student to direct the laser beam through his experiment while I installed a fixed mirror mounted for the laser beam to reach my arm of the experiment. The laser had an interlock system set indicating to the other user by a light signal that an experiment is ongoing and also to prevent unplanned exposure to the laser beam in case of intrusion. My side of the optical bench was fully covered by a black hardboard optical enclosure (Thorlabs) mounted to the table and an enclosure in aluminium was design to be fixed on top of the microscope stage when we do a measurement.

2.2.1 Optical components

The experimental setup for the collection of polarised Raman spectra consists of a compact 785 nm laser (XTRA Toptica Photonics) directed to the back port of an inverted microscope (TI-U Nikon). The light is collected through a 60x oil immersion objective (NA 1.2) and focused on a spectrometer (Andor Shamrock) which records the light on a 1024 x 100 pixel CCD camera (Andor iDUS). The setup design is sketched on the diagram in Figure 2.8.

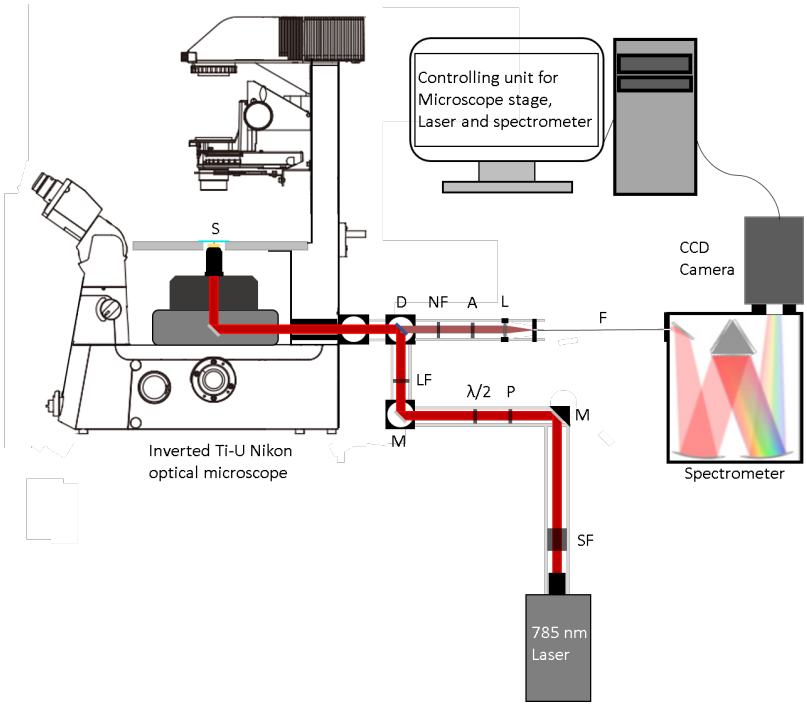


Figure 2.8: Schematic diagram of the Polarised Raman system including spatial filter system (SF), fully reflective mirrors (M), polariser (P), half wave plate ($\lambda/2$), 785 nm laser clean-up filter (LF), 785 nm dichroic beamsplitter (D), sample holder (S), notch filter (NF), analyser (A), focusing lens (L) and fibre optic (F).

The polariser (P) is used to confirm the axis of polarisation of the laser beam, only the laser component parallel to the polariser axis will go through. The half wave plate ($\lambda/2$) allows us to rotate the plane of polarisation of the laser light. In the Raman collection path is placed an analyser (A), a necessary component to tune the direction of collection of the Raman signal compared to the incident polarisation direction. Besides the optical components required to perform polarised Raman measurements, a pinhole system was used for spatial filtering of the laser to obtain the ideal laser beam profile. In laser microscopy, the optimised beam profile is known as the (fundamental) transverse electromagnetic TEM_{00} beam profile. It corresponds to the ideal spatial beam profile of a laser that is very nearly Gaussian and for which the profile does not change as it propagates. In practice, no laser achieves this theoretical performance level [65]. The spatial filtering implemented in our case was intended to tend towards this TEM_{00} beam profile. The spatial filtering consists of focusing the laser beam with an aspheric lens into a pinhole to remove the intensity fluctuations from the Fourier plane

of the object and thus recovering a 'clean' beam after an achromatic lens, as illustrated in Figure 2.9.

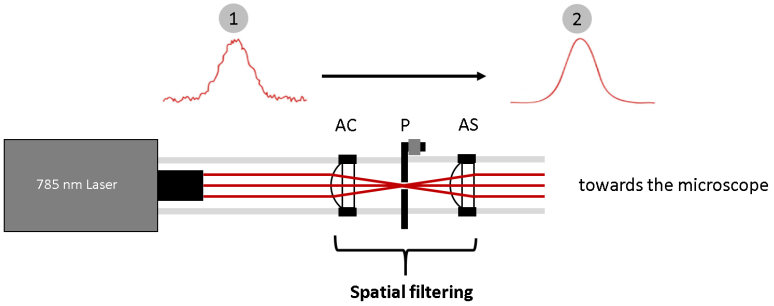


Figure 2.9: Schematic representation of a spatial filter configuration illustrating the removal of intensity fluctuations of the laser beam. The system is composed of an achromatic lens, a 15 μm pinhole and an aspheric lens. The beam from original shape in **(1)** is cleaned to a perfect Gaussian profile shown in **(2)**.

The lenses have matching focal lens (f = 13.86 x 10^{-3} m) and have been chosen to create a compact design. The pinhole size (p_s) satisfies the diffraction limit equation

$$D = \frac{\lambda f}{r} \tag{2.2}$$

with λ the laser wavelength (λ = 785 x 10^{-9} m), and r the input beam radius (r = 1.3 x 10^{-3} m). Using equation 2.2, we find that D is equal to 7.4 μm. Since the pinhole should be at least 30 % larger than the diffraction limit, we chose a pinhole of size p_s = 15 μm.

2.2.2 Calibration

The alignment of the system was confirmed using benzene (C_6H_6) and liquid carbon tetrachloride (CCl_4) samples. They act as standard samples for testing polarised Raman spectrometers since their Raman scattering spectra comprise vibrations for which the intensity will drastically decrease in cross-polarised configuration. Benzene present a very strong symmetric vibration at 992 cm^{-1} from ring stretching [66]. This symmetric band is polarised meaning that its intensity will vanish in cross-polarised configuration. The variation in intensity of this band was measured by rotating the analyser in perpendicular direction to

the laser polarisation. The results are shown in Figure 2.10, the spectra are offset for better legibility.

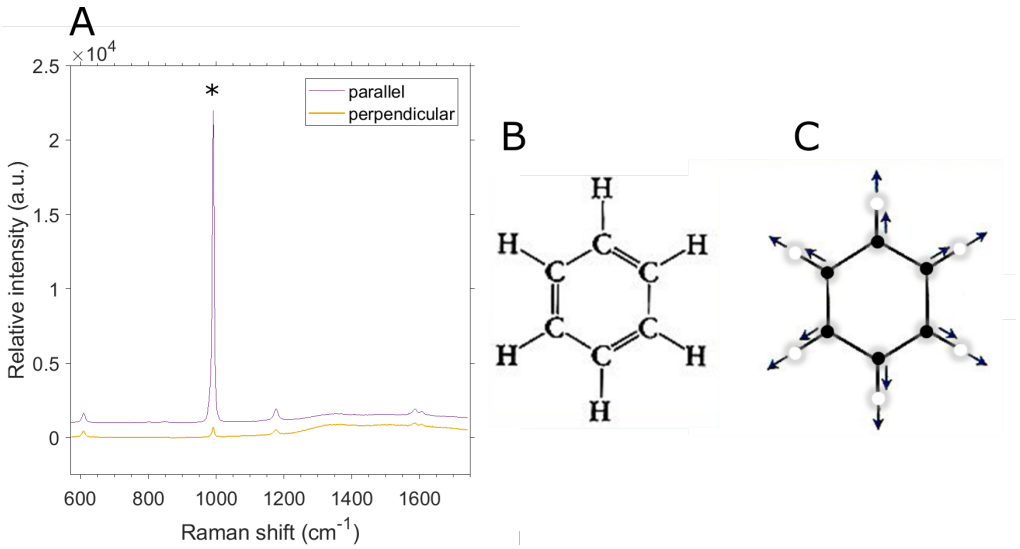


Figure 2.10: **A** Polarised Raman spectra of benzene highlighting (*) the 992 cm^{-1} polarised band. **B** Illustration of the benzene molecule with in **C** arrows indicating the ring stretching $A1_g$ vibration of C_6H_6 at 992 cm^{-1}.

The spectra displayed bands at 606 cm^{-1} from ring deformation and at 1178 cm^{-1} from CH bend whose intensity was unaltered by polarisation as confirmed by Shimanouchi, 1972 [66]. However, we noticed a 96.6 % intensity diminution of the sharp band at 992 cm^{-1} in perpendicular configuration as predicted from theory.

We performed a similar experiment on carbon tetrachloride; it is known to have a pure symmetric vibration at 460 cm^{-1} which will lead to polarised scattered light emission when interacting with polarised incident laser light. By choosing the cross polarised configuration, we were able to make the band ν_1 vibration mode of CCl_4 at 460 cm^{-1} disappear. The experimental results are shown in Figure 2.11.

Typical Raman bands from carbon tetrachloride were observed at 217 cm^{-1}, 314 cm^{-1}, 460 cm^{-1}, 769 cm^{-1} and 795 cm^{-1}. Every band is depolarised except the 460 cm^{-1} band for which the intensity drops to zero in the perpendicular configuration. Both experiments confirmed a correct calibration of the polarised Raman setup developed in this thesis.

We also investigated the laser power stability over time to confirm that the

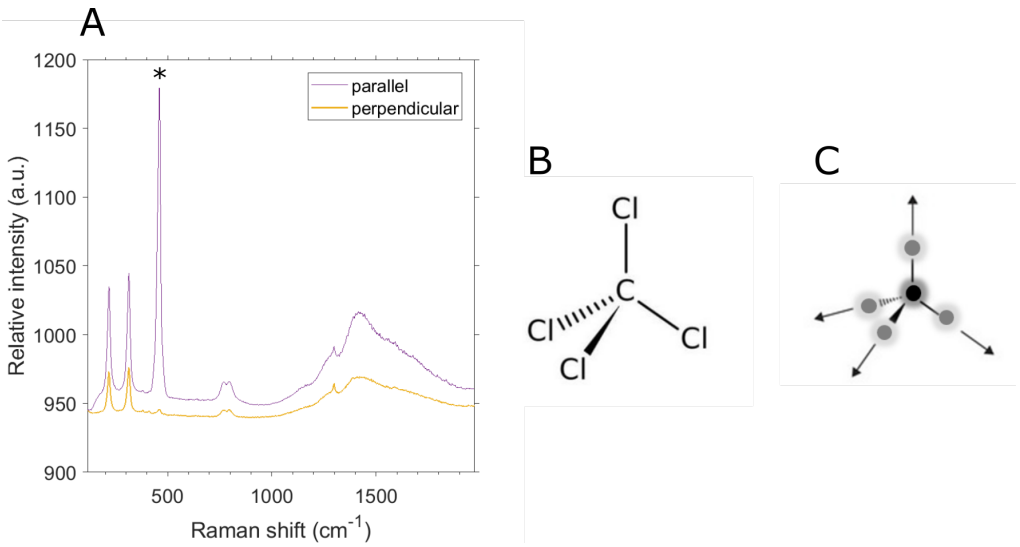


Figure 2.11: **A** Polarised Raman spectra of carbon tetrachloride highlighting (*) the 460 cm^{-1} polarised band. **B** Illustration of the tetrachloromethane molecule with in **C** arrows indicating the ν_1 vibration mode of CCl_4 at 460 cm^{-1}.

measurements are accurate and reproducible. It was important to ensure that the laser power was stable throughout the entire time needed to carry a Raman measurement as well as during a prolonged period, thus no variation could affect the interpretation of the final Raman spectra. The laser power stability curve, recorded over 494 days with a power meter (PM 100D, S121C sensor, Thorlabs), is displayed in Figure 2.12.

The values of the laser power, at a set current (l_{curr} = 1000 mA) at the laser output has the value of 121.9 ± 3.6 mW which means a fluctuation of around 3 %. Concerning the laser power after the filtering system, it was equal to 56.4 ± 3.0 mW expressing a variation of 5.3 %. These results reflect the long term high stability of the laser used in our study implying that no adjustment was required prior to acquiring Raman data.

2.3 Software and spectral analysis

The stage, the spectrometer and the CCD are connected to the computer and controlled using their respective device drivers and dynamic linked libraries. To

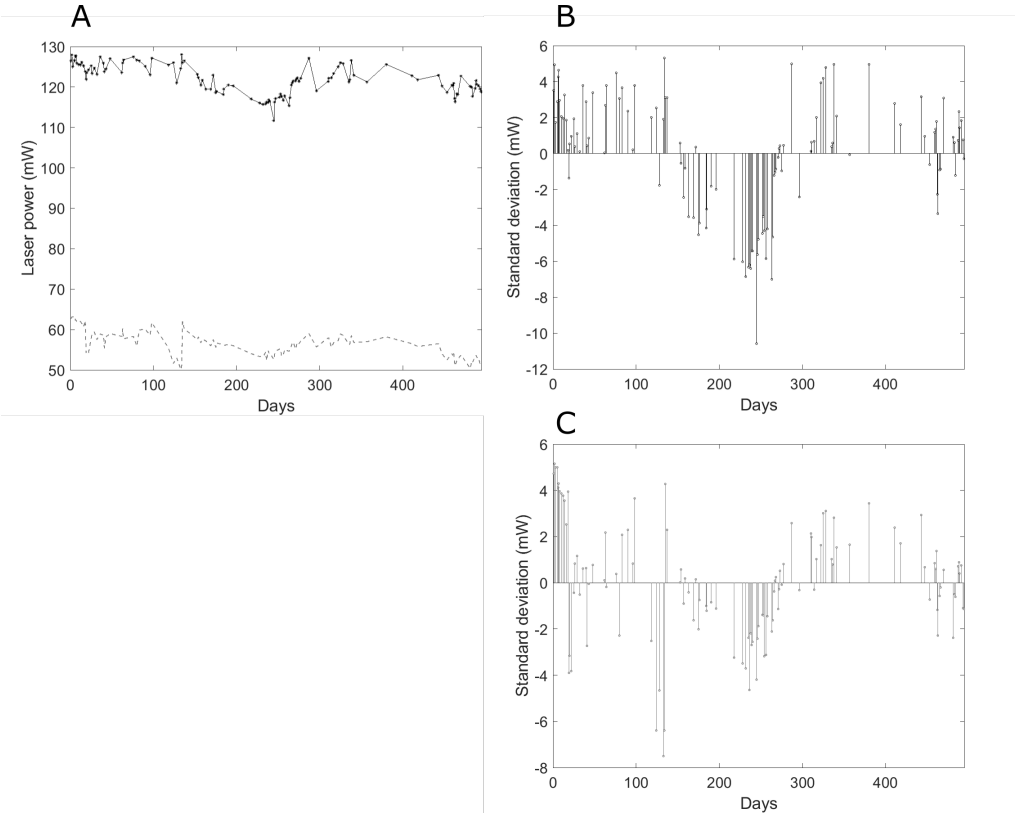


Figure 2.12: **A** Laser power measurement carried between February 2017 and July 2018 at the laser output (dark) and after passing through the spatial filtering system (dotted grey line). The variations of each data point around the mean value at the laser output and after filter are shown respectively in **B** and in **C**.

allow the entire instrument to function synchronically, a customised software was developed within the research group, that was adapted for each project.

LabVIEW

Our experiment setup comprises devices that require to be monitored by the user through a single interface. The LabVIEW software was used as a graphical platform that can control equipment and collect data at once. The programme uses the graphical programming language where an image replaces a function creating a friendly interface.

The interface panel allows the user to:

- permit the communication between all the electrical equipment,
- control the spectrometer by adjusting the cooling temperature of the CCD,
- acquire microscopic images and display the image obtained,

- move the motorized stage and monitor its position in real-time,
- vary the parameters for data acquisition such as scan dimensions, integration time and number of sampling points,
- initiate and stop Raman measurements,
- save the data collected in a format compatible with MATLAB.

A typical Raman experiment starts by recording a microscopic image of the sample from which an area of interest will be selected. The interface allows moving the stage to the desired position within the sample and defines the size of the microscopic image. The program acquires a certain number of sub-images which are tiled in real-time. This ensures the visualisation of the image tile by tile while it is being recorded. Visual control is made possible allowing the user to decide if parameters such as image brightness or contrast might need to be changed or if the acquisition should be paused. Then, the user enters the temperature required for the CCD to operate, the value of the temperature appears on the main panel. No other data acquisition is possible until the chosen temperature is reached. Once the parameters of acquisition are defined by the user, the coordinates of the scanning area are calculated and sent to the stage. The program calculates the memory allocation to ensure enough storage space on the computer. An error message appears if it is insufficient. As soon as the *Start* button is pressed, the stage is moved to the initial position of the scan. During Raman raster scanning, the stage move is automated with the CCD collecting the data until the acquisition is complete or if the *Stop* button is pressed. The data acquired is stored in an array in which each row represents a single Raman spectrum collected at a given position and each column represents the CCD pixel number. The data is stored as a .mat file including the number of photons captured and their corresponding coordinates.

The procedure is similar if we want to acquire a single spectrum. In this case, the Raman point measurement consists of recording Raman spectra over the full pixel range for a defined integration time at a certain and unique position within the sample. Most functions in the program described are run through the MATLAB script node which calls the MATLAB software to execute the actions.

MATLAB

MATLAB software (MathWorks) was used to develop semi-automatic programs for processing the data obtained from each measurement.

The first step in the processing algorithm consists of distinguishing noise events in the Raman scattered spectra. It is necessary to remove any noise disturbance in the collected signal before applying any subtraction or baseline correction to obtain the correct Raman information. The most frequent source of noise is coming from cosmic rays; they are random events caused by cosmic rays hitting the charge-coupled device (CCD) sensor and possible alpha and gamma radiations from the optics [67]. The possible detection of cosmic ray events comes from the high sensitivity of the CCD. In the Raman spectrum, they are recognizable by sharp and intense spikes, they appear in random pixels (usually in only one) with random probability. The cosmic ray removal implies to select peaks from each spectrum with intensities 10 times larger than the standard deviation over the full pixel range. When a cosmic ray is detected, its intensity is replaced by the mean intensity between the left and right support points of the peak.

To minimize the errors when qualifying biomolecules vibrations in Raman studies, we are required to implement a fitting method for background removal. The major source of background noise in Raman spectra for biological studies is the intrinsic fluorescence of organic molecules. The fluorescence background is usually a few orders of magnitude more intense than the transitions involved in Raman scattering. Hence, it can easily overshadow Raman spectra making its analysis impossible. We implemented, throughout this thesis, for fluorescence background removal method a fifth order polynomial fitting as the simplest and most convenient option described in literature [68, 69].

In order to filter and interpret the data from Raman maps, multivariate analysis method were employed [70]. They are used to analyse the data by considering each Raman spectrum as a whole by not considering each peaks individually. Principal component analysis (PCA) and clustering k-means are the most commonly used. PCA brings out the correlation between each pixel while k-means separates the pixels into regions based on likeness of their spectra.

Principal component analysis (PCA)

PCA is a statistical analysis method that performs mathematical decomposition of the spectral data to reduce the dataset dimension to a smaller number of scores, principal components and loadings. This technique finds a new set of axes, called principal components (PCs), to view the data by performing an axis rotation. By doing so, it aligns the new set of axes with the maximal variance in the dataset. The principal components form a new basis of coordinates for the data.The new variables, called scores, show which spectra are different, and the loadings give the contribution of each wavenumber to a particular PC. The PCs are arranged in decreasing order of their contribution to the variance meaning that the first PC contains the maximum variance from the data, the second PC the next highest amount, etc. The principal component represent spectra with variation of the spectral information, they are linear combinations of true spectra with positive and negative intensities. The number of PCs obtained for each map is smaller than the number of spectra in the data. Since PCA reduces the number of variables to describe a system; the scores with significant PCs can be used to represent the spectra.

K-means clustering

To annotate each region from the 2D Raman maps to a specific tissue type, we used the k-means clustering analysis as a data grouping method. This method separates data into clusters by allocating individually each spectrum to the cluster with the nearest mean. In other terms, it generates groups of spectra of similar spectral feature. The aim is to compare the k-means map with the adjacent H&E section. The number of clusters generated to closely fit the k-means map with the H&E section is controlled by the user who wants to avoid over or under grouping. The number of optimal clusters varies between different samples and depends on the type and number of tissue structures annotated from the adjacent H&E section. As a result of this analysis, each tissue type is associated to a single k-means cluster in the k-means map. The k-means maps are pseudo-colour images for which each k-means cluster is assigned to a colour.

Both instruments presented in this chapter were designed, assembled, aligned independently in two different parts of the project at different times. The instruments successfully went through validation procedure validating the axis

calibration, stability (mechanical and thermal), and the resolution reached. We also defined the mathematical analysis used for data interpretation. Both systems perform reliable and reproducible measurements. For the investigation of lung tissues obtained from surgery at Nottingham City Hospital, the Raman spectrometer was used. The Raman mapping results and conclusions of this study are presented in the following chapter (chapter 3). During the second part of the project, the polarised Raman spectrometer was used for polarised Raman mapping for skin tissue characterisation, from skin tissues kept frozen at university. The results will be later presented in chapter 4.

3

Raman mapping studies of fresh human lung tissues

This chapter presents the use of Raman spectroscopy for the imaging and microscopic characterisation of lung tissues ex vivo. We detail the experimental parameters and discuss the results which are correlated with other studies.

3.1 Introduction

Lung cancer is responsible for the most cause of deaths worldwide. Lung cancer has a low five-year survival rate of around 13 % in England [6] representing a major health concern. We described in Chapter 1 section 1.2, the standard procedures used in clinical practice for diagnosis and tumour staging of lung cancer. It is to be noted that early cancer detection via histopathological examination of tissue block resection can take up to two weeks, from the day of bronchoscopy, to provide a diagnosis. This consideration emphasises that one of the main limitations of histopathology is lengthy tissue preparation. This, in addition to the extended time for a patient to access a specialist, could cause late detection leading to higher occurrence of advanced diseases and delay in treatment. Additionally, histology diagnosis relies on morphological characteristics of the tissue, a process that can be biased by subjectivity [3].

Non-invasive optical spectroscopic techniques may provide improved accuracy over the current morphology-based techniques for lung cancer diagnosis. To this day, a limited number of *ex vivo* spectroscopy techniques have been used for the study of lung cancer tissues from which we can cite nuclear magnetic resonance

[71], (auto)fluorescence spectroscopy [72–74], infrared spectroscopy [75, 76] and Raman spectroscopy [36, 77]. This latter technique has been increasingly applied for tissue diagnosis because it is non-destructive so the tissue can be submitted to histopathology after analysis, it has a high molecular specificity [78] and has been demonstrated to achieve high diagnosis accuracy [79–82].

The first study using near-infrared (NIR) Raman spectroscopy for lung cancer tissue study demonstrated the discrimination between healthy and malignant tissue with a specificity and sensitivity of 92 % and 94 % respectively [36]. Recently, for 1064 nm excitation wavelength, spectral differences in the amide I band at 1659 cm^{-1} [77] and bands assigned to collagen 1666 cm^{-1} and 1448 cm^{-1} [35] were used to discriminated between normal and malignant human lung tissues. Although the 1064 nm laser lowered significantly the laser-induced fluorescence background, acquisition times of 5 minutes were required for recording individual Raman spectra.

While these studies confirm the feasibility of Raman spectroscopy for diagnosis of lung cancers, they relied on single spectrum acquisition or a low number of Raman spectra only, which can lead to sampling errors caused by heterogeneity of the lung tissues. Raman mapping of lung tissue has been reported but the measurements were performed on frozen-sectioned healthy lung tissues [83] and congenital pulmonary airway malformation [31] but did not include lung cancer tissues.

In this work, we present the results of the Raman spectral mapping measurements of fresh non-malignant and cancer lung tissue. The aim of the study is to provide a better understanding of the effect of tissue heterogeneity on the Raman signals and investigate whether the spectral differences between healthy and malignant tissue can be used for diagnosis.

3.2 Material and Methods

Tissue collection

Lung tissue samples were obtained from the Nottingham City Hospital. The Nottingham Health Science Biobank (NHSB) provided us with tissues collected after patient consent and ethical approval (NHSB, REC reference 15/NW/0685). All lung specimens were received fresh within 20 min of surgery in the histopathology department. Surgery included left, right, lower, higher or partial lobectomy and thoracoscopy. Tissue sections were stained using Haematoxylin and Eosin (H&E) staining and analysed for diagnostics by a histopathologist.

Group		Sample ID	Tissue type	Description	Date of measurement
GROUP A		Sample A	Tumour		07/10/15
		Sample B	Normal		09/09/15
		Sample C	Normal		08/10/15
GROUP B		Sample D	Tumour		17/08/15
		Sample E	Normal	Interstitial fibrosis and emphysema	08/10/15
		Sample F	Tumour	Adenocarcinoma	12/10/15
		Sample G	Tumour	Adenocarcinoma with hyperchromasia	02/11/15
GROUP C	NORMAL	Sample H	Normal		02/09/15
		Sample I	Normal		11/12/15
		Sample J	Normal		16/06/16
	SCC	Sample K	Tumour	Squamous cell carcinoma	08/09/15
		Sample L	Tumour	Squamous cell carcinoma	17/09/15
		Sample M	Tumour	Squamous cell carcinoma Section 3*	09/10/15
		Sample N	Tumour	Squamous cell carcinoma Section 4*	12/10/15
		Sample O	Tumour	Squamous cell carcinoma	21/01/16
	ADENO	Sample P	Tumour	Adenocarcinoma , various regions, some with necrosis	25/11/15
		Sample Q	Tumour	Adenocarcinoma	08/12/15
		Sample R	Tumour	Lepidic adenocarcinoma	14/12/15

Figure 3.1: Table summarizing the lung tissue samples considered in the study. The * symbol indicates tissue from the same patient but from different section. SCC stands for squamous cell carcinoma, and ADENO stands for adenocarcinoma.

Sample handling and preparation

In the 20 minutes following the lung tissue specimens excision, the fresh tissues were cut in two adjacent blocks: a 'Raman block' (used for Raman analysis)

and a 'Reference block' (used for reference analysis by paraffin-embedded H&E). The 'Raman block' was not immediately Raman scanned, thus it was kept frozen to minimize alteration. Prussian blue ink was used to paint the outer edge of the 'Reference block' to ensure that the tissue orientation was well defined when performing the sectioning for H&E staining. The tissue was then placed in formalin for fixation and then cut into consecutive sections with a microtome. At that stage, the blue ink was a visual marker used to ensure that the first slice of tissue sectioned, named β_1, was coming from the surface of tissue opposite to the dyed region. It guaranteed that β_1 exactly corresponds to the same tissue region to the surface of the 'Raman block'. Each section slide was stained with H&E and analysed by a histopathologist. We used the information obtained from the histopathological analysis from β_1 to select tissue structures of interest for the Raman tissue investigation. Finally, the unstained adjacent section (from the 'Raman block') was defrosted in phosphate-buffered saline (PBS) solution at room temperature and oriented with its surface facing the laser beam on the calcium fluoride (CaF_2) slides. The region of interest from section β_1 defines the corresponding region in the 'Raman block', investigated by Raman spectroscopy. This process is illustrated in Figure 3.2.

Because no chemical treatment was applied to the 'Raman block', it was kept in a state as close as possible as the tissue would be *in vivo*. To avoid the tissue to dry, the tissue was immersed in PBS during the entire Raman measurements.
The method described is a robust way of collecting Raman signals from specific structures of the tissue with the guidance of the H&E images obtained from the 'Reference block'. For most samples, we were able to identify structures from the 'Reference block' and correlate them with the structures in the 'Raman block'. However, for some cases, it was impractical to use the H&E for area selection because the 'Raman block' had a significantly different shape making the comparison difficult, or in certain cases impossible. To address this issue, H&E sections were also prepared from the 'Raman block' after the Raman measurement. This way, the correlation between the Raman scanned location and its corresponding morphological structure was confirmed for measurement accuracy. Figure 3.3

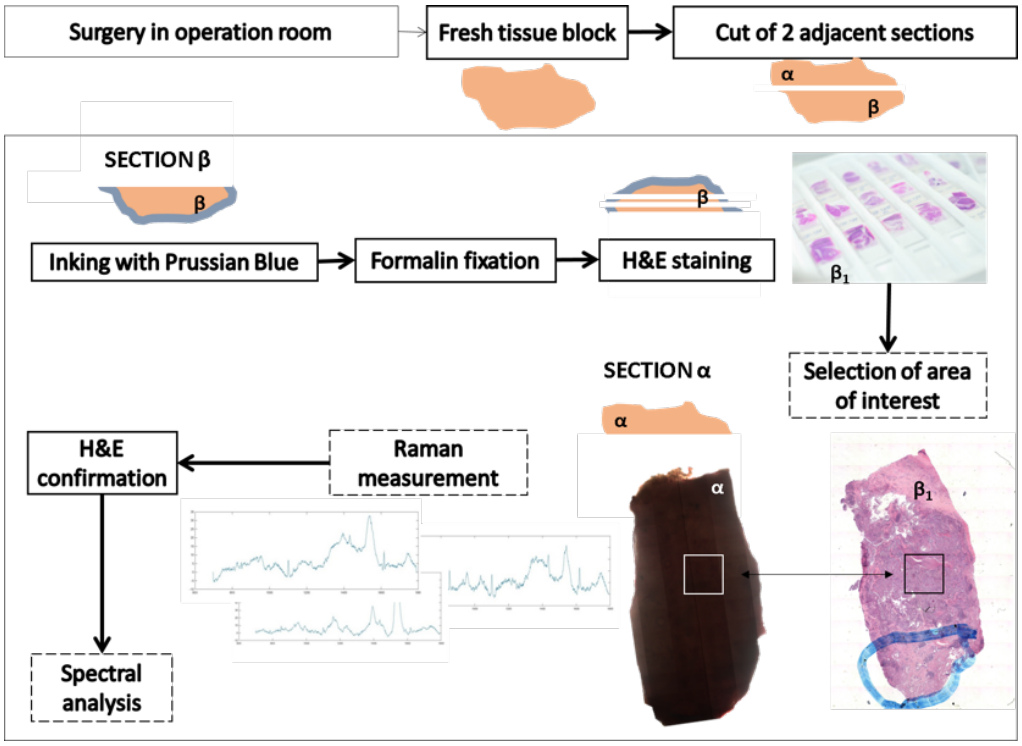


Figure 3.2: Illustration of the procedure of lung tissue handling from tissue resection obtained in operation room to final Raman measurement. The black outlined rectangle boxes indicate the actions of hospital staff (surgeon, histopathologist) while the dashed boxes highlight manipulation in our laboratory. $\alpha$and βrespectively correspond to the 'Raman block' and the 'Reference block'.

is an illustration, for a tissue sample, of the encountered challenge and its solution.

Ex-vivo Raman spectroscopy

The Raman measurements were carried out using the laboratory built and designed Raman microspectrometer described in Chapter 2 section 2.1. The instrument is based on the inverted optical microscope, the Proscan III inverted stage and the deep depletion Andor iDus CCD cooled to -74°C. Lung tissues are placed on the substrate with their surface of interest placed facing down. The substrate consisted of a calcium fluoride coverslip attached to a coverslip holder (Figure 3.4). Cotton was placed on top of the tissue. We added PBS drops to the cotton surface guaranteeing a homogeneous level of hydration of the tissue. It was also necessary to ensure that the tissue surface under investigation was

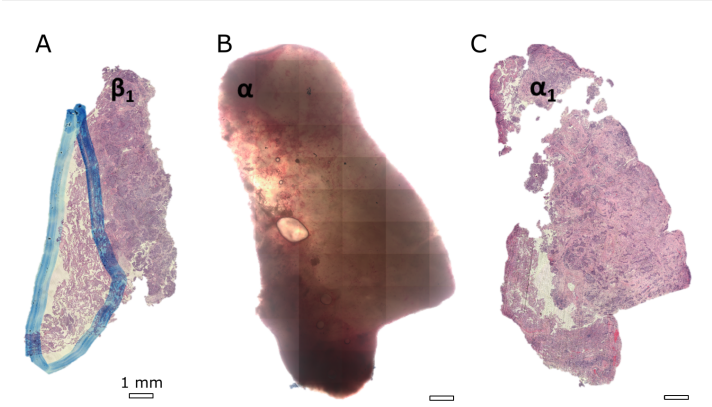


Figure 3.3: **A** shows the H&E stained slide section of the 'Reference block' (β_1) observed under bright-field microscopy. The defrosted tissue from the 'Raman block' (α) seen in **B** has a significantly different appearance than the reference H&E section. It makes it challenging to accurately correlate these two images. For this reason, the 'Raman block' (α) is H&E stained after the Raman measurement to obtain the H&E stained slide section α_1 observed in **C** which is similar to the tissue from 'Raman block' presented in **B**.

pressed against the substrate avoiding tissue regions to not be in contact with the coverslip resulting in being out of focus. We washed the coverslip and the holder with ethanol and PBS after each measurement. The 785 nm laser was focused on the sample through a 60x 0.85 NA immersion oil objective (RiverD International) providing 160 mW laser power at the sample. Raman maps were recorded in mapping mode with 20 μm step size, with an acquisition time of 1 s per spectrum over the spectral range 600 cm^{-1} to 1750 cm^{-1}.

Data processing

The Raman mapping datasets presented consist of 2,500 - 5,000 Raman spectra. Each Raman spectrum was processed for cosmic ray spikes and fluorescence background removal. The fluorescence contribution was subtracted from the spectrum by applying a fifth order polynomial fitting [68, 69]. Then, we performed principal component analysis (PCA) which allows describing the entire set of spectra by a smaller number of parameters [84]. PCA finds the major sources of spectral variance and creates an orthogonal base of a new system of coordinates called the PC scores. We can present the result in binary scatter plot of PC-i vs PC-j with i, j = 0, 1, 2, ... useful for separating the data in clouds

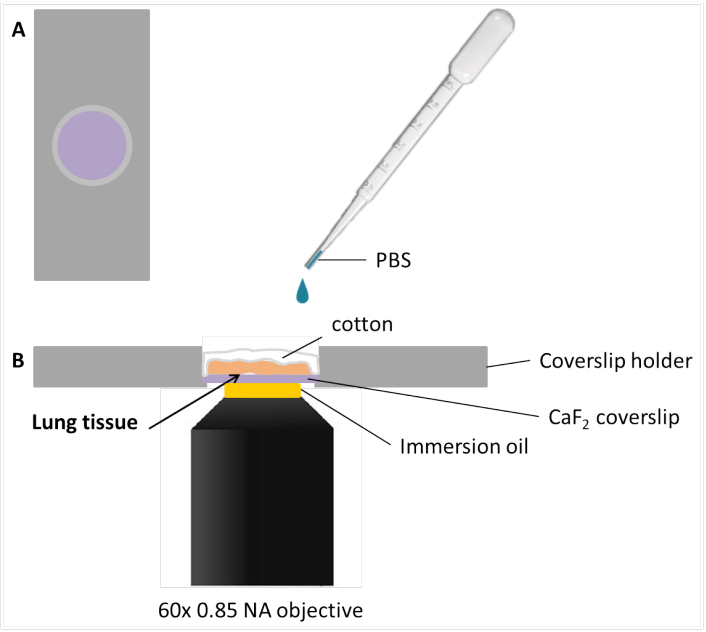


Figure 3.4: Schematic representation of the inverted setup for the measurement of lung tissues showing the immersion oil objective, the immersion oil, the sample coverslip and holder, the lung sample, the cotton pressing the sample and the pipette loaded with PBS **B**. In **A** is a top-view representation of the custom coverslip holder (in grey) and the calcium fluoride coverslip (in light purple).

where a diagnostic line based on a visual examination can be drawn. Besides, the PC scores are used as an input for k-means cluster analysis. By assigning a color to each k -means cluster,pseudo-color Raman images were generated. The k-means clustering image was helpful to correlate the Raman spectral maps with the adjacent H&E images.

3.3 Results

Measurement parameters

The acquisition time for a Raman mapping measurement depends on parameters such as the width (w) and length (l) of the area studied, the acquisition time (t) of each spectrum per pixel, the step size (s) of the scan and various experimental restrictions including the detector sensitivity (lower temperature, higher

load resistance). We optimised the Raman instrument by selecting the suitable components, such as the CCD camera, spectrometer, laser and objective [85].

Data acquisition time

The acquisition time is limited by the spectra quality required for the Raman signal to contain the required spectral information to allow extraction of the desired information. Because the Raman signals are quite weak in general, compared to fluorescence or absorption, the signal-to-noise ratio (SNR) is required to be the highest possible. For example, previous studies suggested that SNR greater than 2 is sufficient to obtain meaningful results in the Raman quantitative analysis [86, 87]. The SNR for a specific Raman band (y) is defined as

$$SNR = \frac{\bar{S}}{\sigma_y} \tag{3.1}$$

Where $\bar{S}$ represent the peak height and σ_{y} the standard deviation of the peak height [26]. Longer acquisition times and higher input laser power lead to a higher signal-to-noise ratio. The laser power used in this was chosen to avoid tissue damage at 785 nm, it is therefore not a possible option to increase its power.

Concerning the acquisition time (t), 0.5 s, 1 s, 2 s, and 5 s were tested on biological tissues, and there was an increase of the SNR with the integration time. The one second collection time provided the required Raman SNR values (SNR $>$ 2) while staying as short as possible for a realistic measurement time frame. As an illustration, the histogram in Figure 3.5 revealed the value of SNR for each spectrum obtained during a typical Raman mapping containing 2500 spectra. At least 68 % of the spectra have a signal to noise ratio between 2.08 and 3.52, it indicates that only around 15 % of the spectra with poor signal to noise ratio (SNR $<$ 2).

We should also mention that the spectra are computationally processed after their acquisition to improve the SNR by signal averaging when grouping spectra from PCA analysis into clusters. From the PCA analysis procedure, spectra from the same cluster are combined. Each spectrum is affected by some sort of noise. The source and properties of instrumentation noise in Raman signals are classified

into two groups. The first group is the so-called Gaussian or white noise; it is when the noise in the signal fluctuates randomly about a zero mean with its statistics independent of time. The second group represents the noise caused by random external events creating huge absurd spikes [88]. These spikes are removed from the signal by the CR removal procedure. However, when the signal is contaminated by white noise, we expect that by signal averaging, the noise adds destructively when the signal builds up. It exploits the idea that when doing a measurement multiple times, the signal part will tend to accumulate while the irregular noise will tend to cancel itself, leading to the fact that signal averaging increases the signal-to-noise ratio. Indeed, the standard deviation of the mean of N measurements is smaller by a factor of $\sqrt{N}$ compared to the standard deviation of a single measurement.

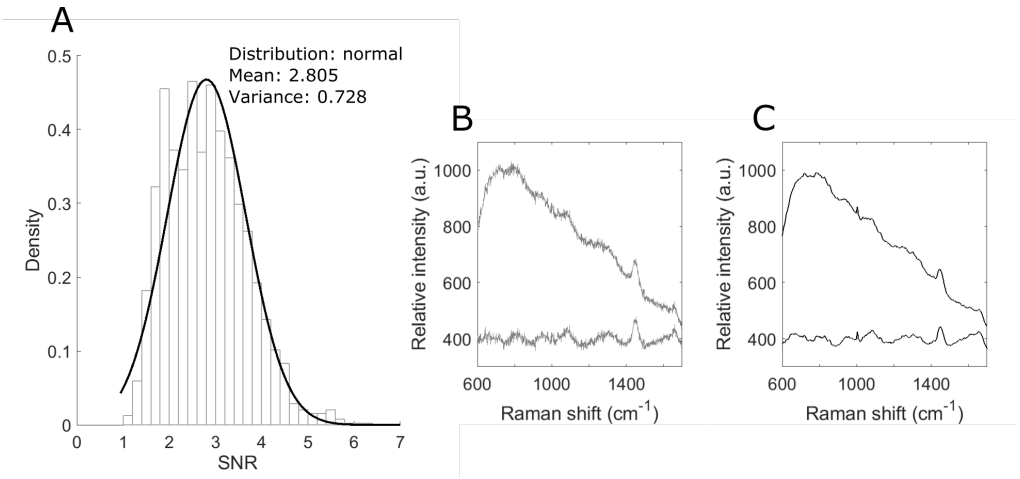


Figure 3.5: **A** Histogram representing value of signal to noise ratio (SNR) for each spectrum recorded with, in black, a Gaussian line fit. Raw and fifth-order fitted spectrum for SNR = 3.53 in **B** and corresponding principal component cluster in **C**.

Area

For correlation between the H&E image and the Raman maps, it is convenient to have square Raman mappings or maps of at least 1 mm length. For this reason, for most measurements we choose the width w and length l of the area to both equal 1 mm, and in some cases, l = 1 mm and w = 2 mm. Accordingly, in all cases, w and l satisfies the condition $w \cdot l \leq 2\ mm^2$.

Since the acquisition time for each spectrum (t = 1 s) and the map dimenstions ($w \cdot l = 2\ mm^2$ to the maximum) were set, the next parameter we considered was the step size.

Step size

To ensure that no significant changes in the molecular composition of tissue occurred during the Raman mapping, the measurement time was limited to 2 hours. The step size of the measurement defines the number of spectra acquired for an array of points. It is constant in the x and y-directions and need to be relevant for biological studies.

The total time (T_t) of measurement of a Raman map is defined by equation 3.2.

$$T_t = \frac{w \cdot l \cdot t}{s \cdot s} \tag{3.2}$$

The allowed total time for the Raman measurement needs to satisfy $T_t <$ 90 min corresponding to the two hours limit and the time needed to prepare the experiment.

It implies that equation 3.2 becomes

$$s \cdot s = \frac{w \cdot l \cdot t}{T_t} = \frac{2(mm^2) \cdot 1(s)}{5400(s)} = 3.70 \cdot 10^{-4}\, mm^2 \tag{3.3}$$

Thus giving

$$s \cdot s = s^2 = 3.70 \cdot 10^{-4} \Rightarrow s = \sqrt{3.70 \cdot 10^{-4}} = 0.019\, mm \tag{3.4}$$

The step size was then defined as s = 20 μm, which is suitable for our purpose, and on a similar scale to previous Raman mapping studies on lung tissues [31].

Shrinkage effects

The complexity in correlating the Raman map and its corresponding H&E region could also come from the change of shape of the tissue during formalin fixation, tissue processing and exposure to laser beam *ex vivo* happening during the Raman measurement. Tissue fixation was performed to stop the tissue degradation that takes place as soon as a biological tissue is deprived of its blood supply. Fixation preserves the tissue as closely as possible to its living state which is necessary for accurate tissue diagnosis. Measurements of tissue dimensions were taken with a ruler from the microscopic images from fresh tissues placed on the Raman slide

and from the tissue after tissue processing. Tissue shrinkage for width and height were calculated by equation 3.5

$$\begin{aligned} \%shrinkage\ (w) &= 100\ - \frac{w_f \cdot 100}{w_i} \\ \%shrinkage\ (h) &= 100\ - \frac{h_f \cdot 100}{h_i} \end{aligned} \tag{3.5}$$

A positive shrinkage percentage in a direction indicates a shrinking of the tissue in this direction while a negative express the expansion of the tissue after tissue processing. The results of tissue-shrinkage evaluation for three samples are summarized in Figure 3.6.

We noticed a tissue shrinkage in the width direction of the tissue ranging between 15 % and 31 %. There was a systematic shrinkage of the tissue in this direction when exposed to the H&E staining routine. However, the behaviour was different in the height direction where 22 % of shrinkage was observed for one sample only. For the two others samples, one did not present shrinkage in this direction and on the contrary, the last tissue displayed tissue relaxation in this direction. The effect of histologic processing causing tissue shrinkage has been described in the literature for different types of tissue. The existence of shrinkage has been linked to variables such as age and gender of the patient, site from where the tissue was excised and its constitution, and parameters of formalin fixation process such as fixation duration [89]. It was reported that formalin fixation may cause lung tumour shrinkage [90], the study reported that after formalin fixation the shrinkage rate, defined by the cross-sectional diameter, varies between 0 % and approximately 29 %. Despite, the measurement method being different compared to our study, our values fall in the shrinking rate range mentioned.

To limit patient variability, we ran measurements on three samples corresponding to three sections cut from the same lung tissue block (Figure 3.6). Although we would expect the samples to present similar behaviours, we noticed that the shrinkage rate shows disparate results. Manual handling factors such as fixation duration and section cutting method influence the shrinking. It was shown that the prolonged formalin fixation causes structural changes capable of inducing tumour size variations [90], while the tissue cutting method involving

A

SAMPLE number	Fresh state (in mm)	After processing (in mm)	Variation (in mm)	Shrinkage
	Width (w_i) Height (h_i)	Width (w_f) Height (h_f)	In width (Δw) In height (Δh)	%shrinkage(w) %shrinkage(h)
1	w_i = 4.7 h_i = 16.9	w_f = 3.7 h_f = 13.2	Δw = 1.0 Δh = 3.8	**21.3 %** **21.9 %**
2	w_i = 4.4 h_i = 14.8	w_f = 3.0 h_f = 14.8	Δw = 1.4 Δh = 0	**31.2 %** **0 %**
3	w_i = 4.7 h_i = 16.1	w_f = 4.0 h_f = 16.8	Δw = 0.7 Δh = - 0.7	**14.9 %** **- 4.3 %**

B

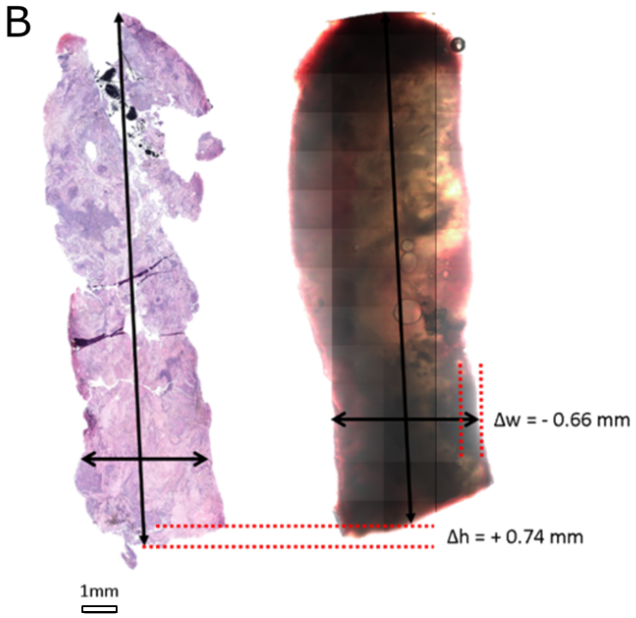


Figure 3.6: **A** Table representing the measurement of tissue shrinkage of three lung samples numbered as 1, 2 and 3 performed in fresh state and after tissue processing. The width $w_{i,f}$ and height $h_{i,f}$ are measured in mm, the variation variable is defined as $\Delta w = w_i - w_f$ respectively $\Delta h = h_i - h_f$. The shrinkage is expressed in percentage. w, h, Δw and Δh have an uncertainty of $\pm$ 1 mm from the random error in reading distances. In **B** is shown the microscopic images of the third sample from which are evaluated the mentioned dimensions with on the right the lung tissue sample on the quartz slide and on the left its corresponding H&E stained slide.

block surface tissue inclination during sectioning will also affect the tissue size measurement [91]. From all these considerations, we conclude that further research is needed to provide evidence on which factors induce lung tissue shrinkage. In this study, we did not implement a correction factor, for instance, described in another study as the shrinkage factor [91]. In addition, we briefly

also mention that we did not notice any change in Raman spectra signal quality (SNR) between fresh tissue and fresh tissue frozen for few days previous the measurement. Spectra from four tissues were compared and visually assessed. We did not notice any impact on the spectra due to the freezing procedure; we concluded that freezing does not affect spectral quality (Figure 3.7). The fresh tissues were measured as soon as they were received from the surgery, while the frozen tissue where kept in the freezer for a couple of days. When we were ready to record their Raman spectra, we immersed the tissue samples in PBS in a glass Petri dish for them to reach ambient temperature. Only then we ran the Raman spectral measurement.

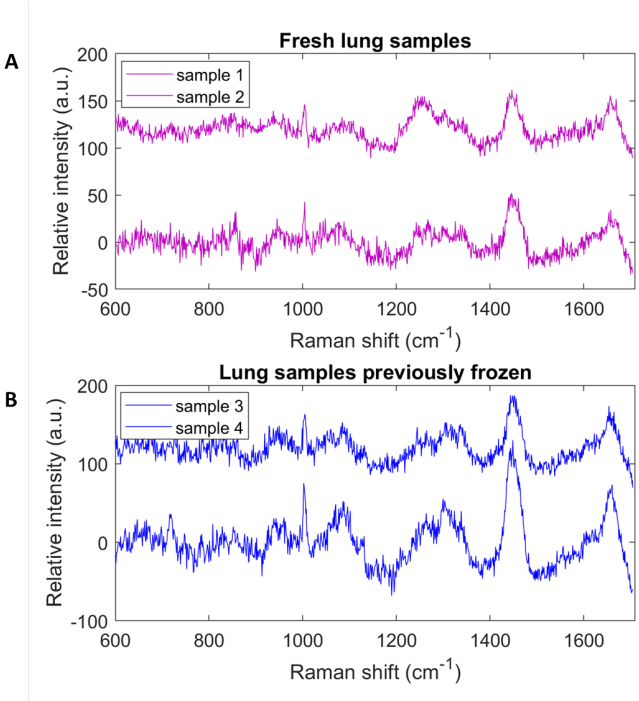


Figure 3.7: Comparison of Raman spectrum recorded for 1 s integration time of **A** two fresh tissue samples (sample 1 tumour SCC, sample 2 healthy) and **B** two tissue samples (sample 3 and sample 4 tumour SCC) which were kept frozen for a couple of days before their Raman measurement.

We investigated a total of 18 lung tissue samples which results can be classified into three large groups: A, B and C. The group A (n=3 samples) includes the overheated samples, the tissues in group B (n=4 samples) are the ones for which the Raman signal is hidden in a broad intense fluorescence background and finally the group C (n=11 samples) is the most populated group containing the tissues with Raman information.

Group A: tissues with laser damage

We observed in three cases, extreme tissue deterioration due to laser exposure (Figure 3.8). As part of the study, we proceeded to Raman mapping measurements in the usual conditions on the tissues shown on the bright-field images in Figure 3.8 **B**, **F** and **I**. At the end of the measurement, we could clearly visually notice large dark indication of burned tissue. The resulting H&E image obtained from the measured tissue corroborate the observations. In the first case, the measurement had induced a breakage in the tissue as seen in Figure 3.8 **C**, for the second tissue, a burn mark was distinctly identifiable in Figure 3.8 **G**, and for the third tissue, a hole was created where the laser beam was hitting the tissue as shown in Figure 3.8 **J**.

The raw spectra collected from the sample displayed at the top of the Figure 3.8 were spread over a large range of intensity and 28.8 % of the spectra reach the maximum charge transfer capacity of the CCD equal to 6.5×10^4 photons count. Concerning the sample presented in the middle, the first row of its Raman map (the first 50 spectra) showed Raman peaks while for the following 120 spectra, the CCD detector saturation was once again reached due to the overheated tissue sample, indicated by the horizontal line observed at 6.5×10^4 a.u. A similar behaviour was observed for the two other samples.
These three tissues samples were three different tissue types, the top one was a tissue with tumour cells whereas the other two were both healthy lung tissues (with emphysema). We could not establish a correlation between the tissues to explain why the tissues underwent thermal damage during the Raman measurement but the laser-induced tissue damage could be due to extremely focused laser light. The laser power density responsible of this burning event can be reduced by changing the numerical aperture (NA) of the objective to enlarge the laser spot, by lowering the laser power or reducing the laser exposure time. However, since this phenomenon was observed for only 15 % of the samples (in a random manner), we did not want to proceed to any experimental changes.

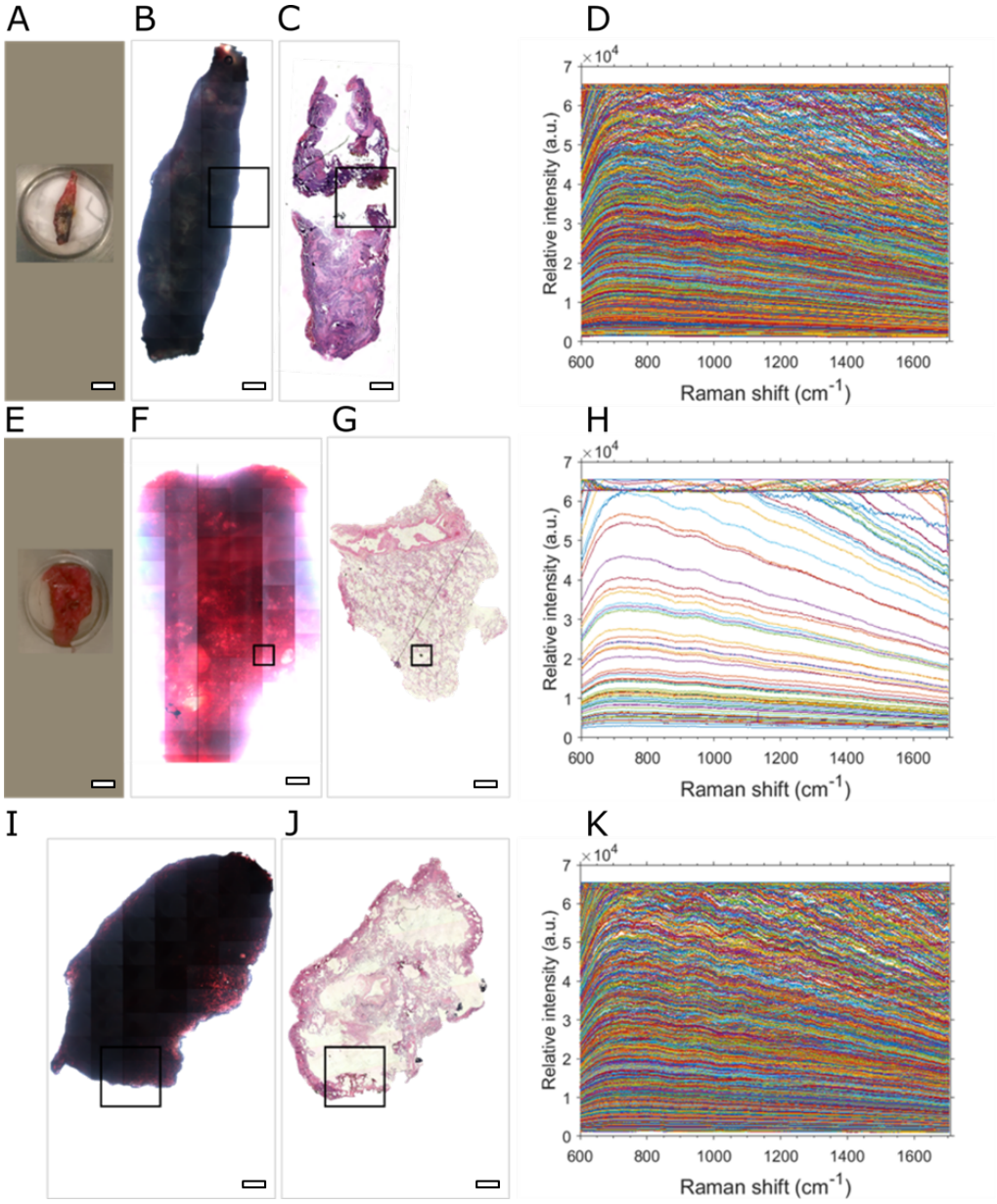


Figure 3.8: Observation of tissue deterioration.
Samples in **B**, **F** and **I** are Raman measured in the region outlined in black. The spectra obtained are presented respectively in **E**, **H** and **K**. In **A** and **D** are the pictures of the two first samples, placed on their sample holder, after the Raman measurement was performed (bar scales indicate 4.4 mm). Each tissue corresponding H&E image is shown in **C**, **G** and **J**. Microscopic images (4x objective) and H&E images have their scale bars equal to 1 mm.

Group B: tissues with strong autofluorescence

For 4 out of the 18 tissue samples investigated, the Raman bands were masked under a strong laser-induced fluorescence background, at all locations in the

Raman maps. Samples presenting this behaviour were both non-malignant and malignant lung tissues. The non-malignant and adenocarcinoma tissues displayed in Figure 3.9 serve as illustration. The H&E of the tissue in Figure 3.9 **B** indicated that the analysed region had emphysema and interstitial fibrosis. For this tissue, the maximum fluorescence intensity varied between 4,000 to 14,000 counts. On the other side, the adenocarcinoma sample consisted of nests of tumour cells with hyperchromasia. The maximum intensity of the fluorescence emission recorded ranged from 20,000 counts to higher than 65,000 counts which is the maximum pixel well of the spectrometer CCD.

For this group, the Raman spectral mapping measurements had a high variance in the intensity of fluorescence emission, even for regions of the sample considered structurally homogenous after histopathology evaluation. No structure was identified in the tissue to justify the source of the laser-induced fluorescence emission.

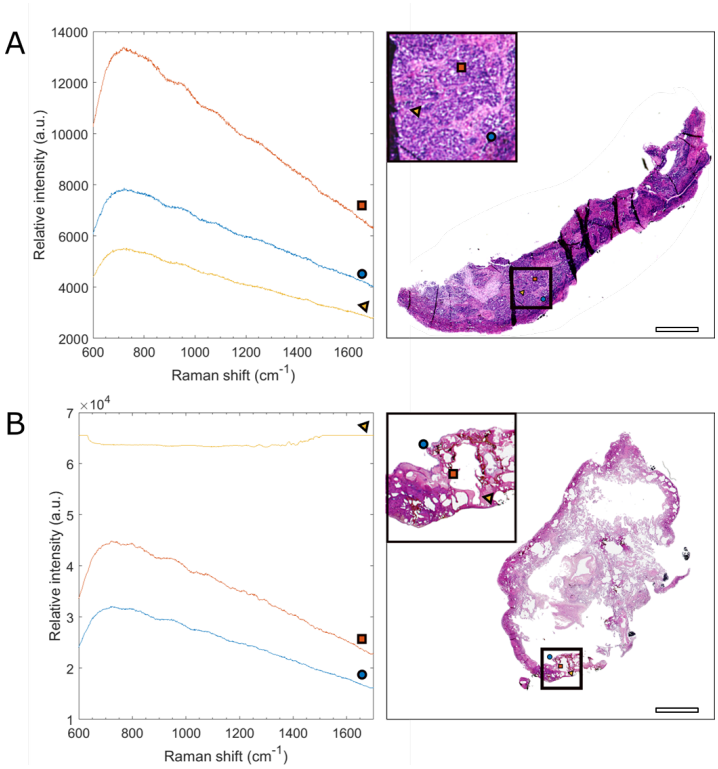


Figure 3.9: Spectral mapping of lung tissues that generated strong laser-induced fluorescence emission that fully swamped the Raman bands. Left panels present typical spectra at three selected positions, indicated by a circle, a square and a triangle, represented in the corresponding H&E images (right panels). Insets represent the tissue areas analysed. The tissue in **A** is a sample with adenocarcinoma and in **B** is a non-malignant lung tissue. Scale bar indicates 1000 microns.

Group C: tissues with Raman spectral information

This group consisted of the remaining 11 samples for which well-defined Raman bands were detected in the collected spectra, despite the presence of a relatively strong fluorescence background, in agreement with previous reports [36, 38]. These samples included 3 non-malignant lung tissues, 5 tissues presenting squamous cell carcinoma, and 3 tissues with adenocarcinoma (details in Figure 3.1).

In total, 17,700 Raman spectra were collected from non-malignant lung tissue samples from three distinct patients. Figure 3.10 presents typical examples of Raman spectral maps and centroid Raman spectra after k-means analysis of spectral datasets.

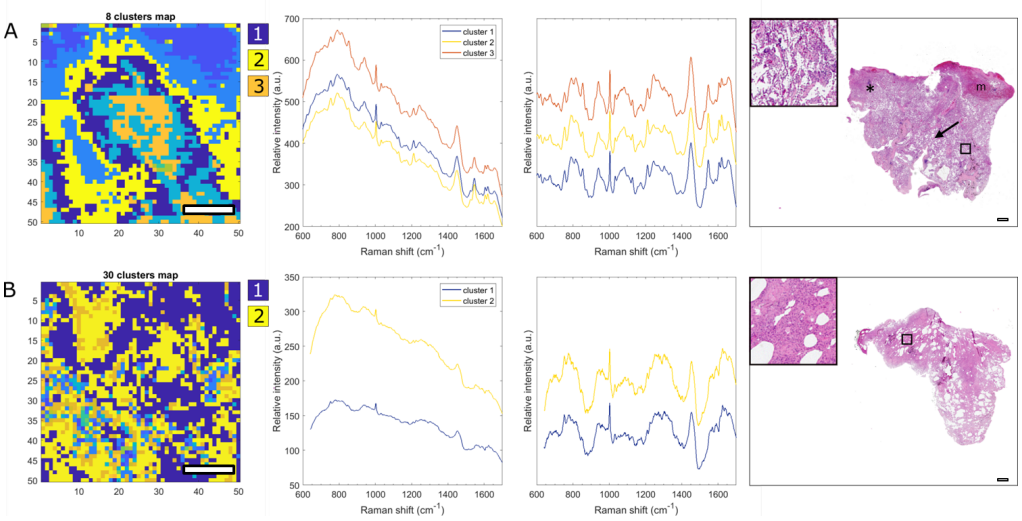


Figure 3.10: Spectral mapping of non-malignant lung tissue (two typical samples included **A** and **B**). Left panels present spectral maps obtained by k-means clustering from which were selected spectra corresponding to the clusters containing Raman bands of highest intensity (scale bars indicate 250 microns). The spectra are presented in the two middle panels prior and after pre-processing while the corresponding H&E images are displayed in the right panels with insets of the tissue areas analysed (scale bars indicate 1000 microns). The region with fibrosis is highlighted by an arrow, a region with inflamed cells is highlighted by an asterisk while the 'm' shows a region with macrophages.

The spectra measured from the non-malignant tissues contain Raman bands that can be assigned to all major groups of biomolecules found in lung tissue [31,

35, 36, 38, 92]. Raman bands associated to molecular vibrations in proteins including protein backbone C-C vibrations (935 cm^{-1}), amide I (1655 cm^{-1}), amide III (1266 - 1360 cm^{-1}). We observed intense Raman bands from molecular vibrations of phospholipids at 1078 cm^{-1}, 1302 cm^{-1} and 1445 cm^{-1} and Raman bands specific to amino acids, such as phenylalanine (1003 cm^{-1}), proline and valine (935 cm^{-1}), tryptophan (1266 cm^{-1}, 1618 cm^{-1}). The Raman spectra also contained bands from nucleic acids at 788 cm^{-1}, 1098 cm^{-1}, 1322 cm^{-1} and 1335 cm^{-1}. Additionally, contributions from blood were identified by the Raman bands at 1370 cm^{-1}, 1577 cm^{-1} and 1620 cm^{-1} which were also identified by Krafft et al. and assigned to haemoglobin [31]. Compared to previous studies, we did not observe the Raman band at 1123 cm^{-1} from proteins reported by Huang et al. [36] instead we detected a band at 1128 cm^{-1}. Moreover, additional bands were observed at 1546 cm^{-1} and 1605 cm^{-1} attributed to tryptophan and phenylalanine respectively, in the case of the sample A [36].

The pseudo-colour images obtained by k-means clustering analysis of the Raman measurements confirmed the heterogeneity of the samples at the molecular level. The sample A includes fibrosis, macrophage and inflamed cells. The analysis of the centroid Raman spectra indicated important differences related to both Raman bands and fluorescence emission. Variations in protein content were indicated by differences in band intensities at 855 cm^{-1}, 935 cm^{-1}, 1303 cm^{-1}, 1252 cm^{-1} and collagen bands at 935 cm^{-1} and 1252 cm^{-1} [92]. The sample B, which consists of denser collagen than the sample A, present emphysema, chronic inflammation and interstitial fibrosis. While the intensity of the fluorescence emission varied by 50 % within the three major clusters, the main differences in the Raman spectra were observed principally at 1078 cm^{-1}, 1128 cm^{-1} and 1208 cm^{-1} which can be associated to phospholipids, proteins and tryptophan and phenylalanine respectively [36, 92].

Then, five tissue samples from four patients with squamous cell carcinoma were analysed (total of 16,300 Raman spectra). Two typical examples of Raman maps analysed by k-means clustering are presented in Figure 3.11. Overall, the spectra of squamous cell carcinoma showed Raman bands also present in normal lung tissue, such as bands at 823 cm^{-1}, 876 cm^{-1}, 935 cm^{-1}, 1004 cm^{-1}, 1123 cm^{-1},

1152 cm^{-1}, 1265 cm^{-1}, 1302 cm^{-1}, 1335 cm^{-1}, 1445 cm^{-1}, 1552 cm^{-1}, 1618 cm^{-1} and 1655 cm^{-1} [36].

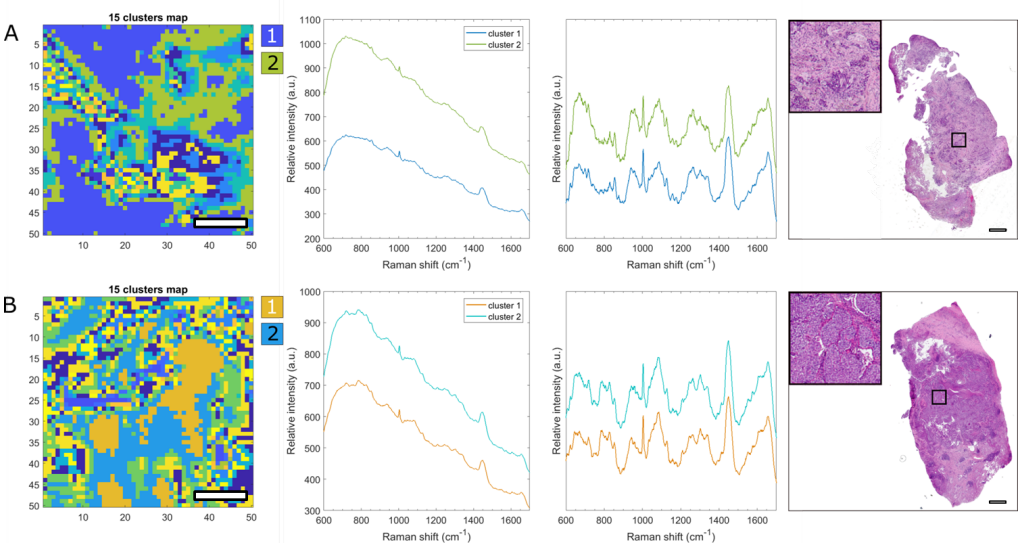


Figure 3.11: Raman measurement of lung tumour tissues **A** and **B** comprising squamous cell carcinoma cells (SCC). The area highlighted in the microscopic H&E image (right panel) has its k-means cluster image represented in the left panel (scale bars indicate 250 microns). The average unprocessed Raman spectra alongside the normalised spectra from the mentioned (1 and 2 respectively) clusters are displayed in the two middle panels (scale bars indicate 1000 microns).

The microscopic molecular and structural heterogeneity of the lung tissue samples was well captured by the pseudo-colour images obtained by k-means clustering. The histology H&E images indicated that the samples present poorly differentiated squamous cells. Also, they had regions of numbers of tumour cells in stroma which were well discriminated in the Raman spectral maps. For the sample in Figure 3.11 **A**, the centroid spectrum from the cluster 2 correlate with the H&E to a region of connective tissue, distinguished by stroma fibres mainly and tumour cells. On the other hand, cluster 1 represents a region of highly concentrated cells, grouped together, they are described as lymphoid follicles in tumour stroma with germinal centres. The region from cluster 2 is richer in collagen content compared to the other regions in the tissue. The corresponding spectra exhibited a higher level of fluorescence intensity. The centroid spectrum from cluster 1 was characterised by significantly lower fluorescent intensity (2-fold lower). The sample region in Figure 3.11 **B** is an area highly concentrated in tumour cells with thin stroma walls. The clusters orange and cyan represent regions with a high

content of tumour cells while the yellow and green cluster represent regions thin stroma. The spectra collected from the last clusters mentioned present level of fluorescence background reaching 2760 a.u. counts for the peak at 750 cm^{-1} (data not presented). Once more, the intense fluorescence background was associated with regions of lower content of tumour cells. The orange and cyan clusters, with higher number of cells, led to higher intensity of the Raman bands assigned to nucleic acids at 788 cm^{-1} (phosphodiester bond O-P-O stretching vibration) and 1098 cm^{-1} (phosphodioxy PO_{2-} group) [93].

Lastly, the specificities of lepidic adenocarcinoma and adenocarcinoma with necrosis were characterised by Raman mapping measurements. A total of 18,825 Raman spectra were collected from adenocarcinoma tissues from three patients. Figure 3.12 presents two typical examples.

Sample **A** in Figure 3.12 present a lepidic pattern of cancerous gland. Based on the H&E image, the region measured by Raman spectroscopy was structurally homogeneous. The two major clusters (cluster 1 and 2), from the pseudo-coloured k-means image, cover 71 % of the area and their corresponding centroid spectra contained well-defined Raman bands. These clusters agree with the pattern observed in the H&E image, they represent the tumour cells lining the alveolar walls. Overall, the centroid spectra of tissue **A** displayed similar bands and fluorescence background than observed for SCC and normal tissue. Yet, we can discern differences. The main difference was the band at 1209 cm^{-1} band, assigned tentatively to tryptophan and phenylalanine [94], which was more pronounced compared to SCC. Also, we observed differences in the shape of both bands at 1302 cm^{-1} and 1335 cm^{-1}. Finally, the Raman band assigned to tryptophan at 1618 cm^{-1} was more intense for the adenocarcinoma sample A compared to the SCC samples.
The second tissue presented in Figure 3.12 had a highly heterogeneous structure. Four regions were investigated by Raman mapping. These were regions of early stage necrosis (Figure 3.12**B**), late-stage necrosis (Figure 3.12**C**), a mixed region of viable cells, necrosis and stroma (Figure 3.12**D**), and a region of stroma with viable tumour cells grouped in a gland shape (Figure 3.12**E**). The low level of structural heterogeneity of the region of early necrosis was confirmed by the

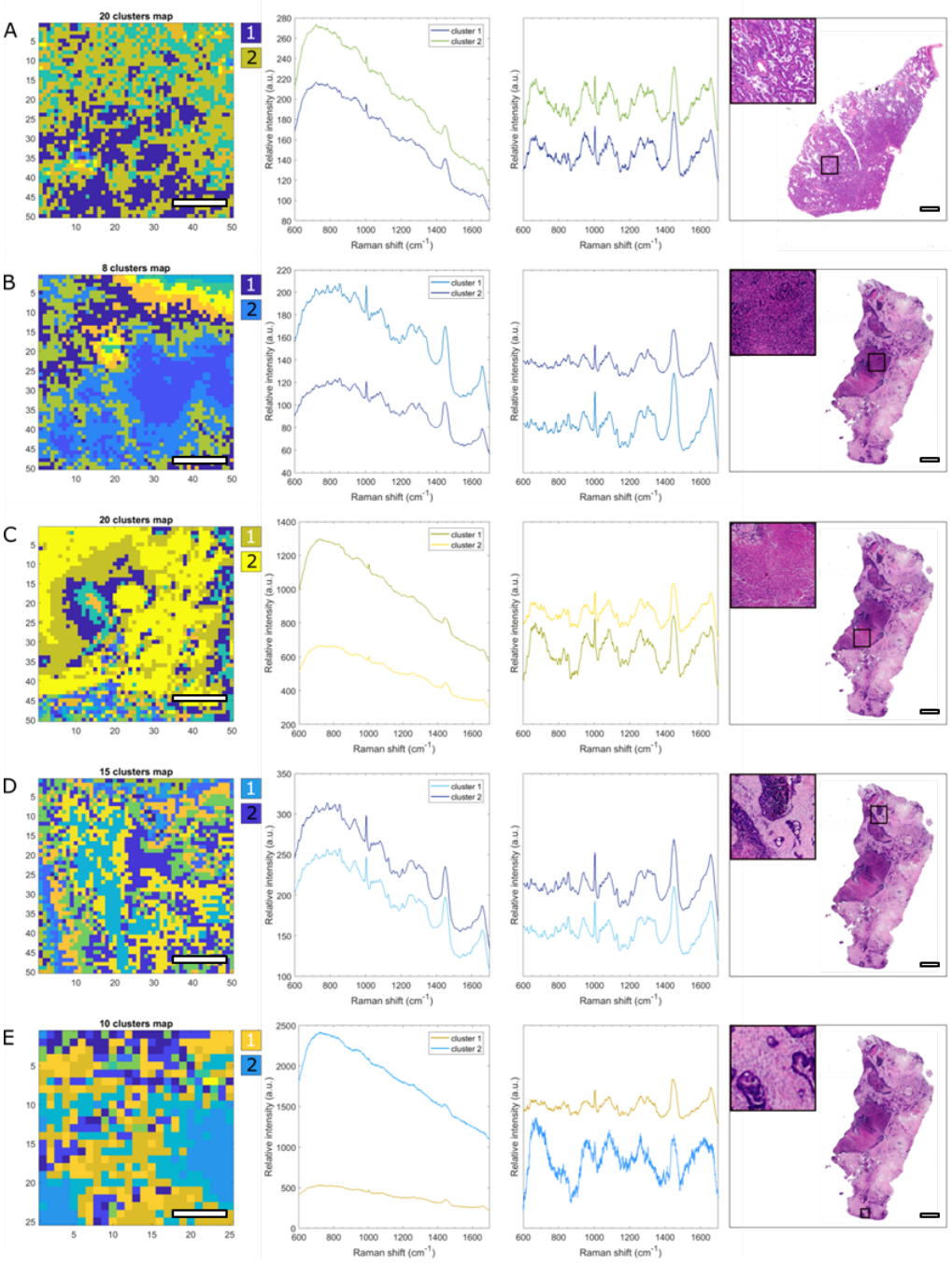


Figure 3.12: Raman measurement of lung adenocarcinoma tissues. The tissue in **A** present lepidic pattern adenocarcinoma. **B**, **C**, **D** and **E** represent various tissues structures of the same tumour tissue presented in the right panels. The regions of early stage necrosis **B**, late stage necrosis **C**, of viable cells, necrosis and stroma **D**, and of stroma with viable tumour cells grouped in gland shape **E** have their k-means based image displayed in left panels (scale bars indicate 250 microns) with corresponding average Raman spectra plotted (scale bars indicate 1000 microns).

overall similar Raman spectral characteristics of the centroid spectra from the 8 clusters. However, the spectra of this region showed the lowest levels of laser-

induced fluorescence compared to all other samples investigated in this study. 20 clusters were then used to analyse the Raman dataset acquire from the late necrosis region. The 4 clusters, accounting for 87 % of the Raman map, presented clearly discernible Raman bands and fluorescence level varying between 700 and 3,500 counts. The spectra contained a peak at 1415 cm^{-1} not assigned in the literature and not observed in any other adenocarcinoma tissue structure. We also observed band shape differences in the region 1550 - 1700 cm^{-1}. This region had more intense marked peaks for porphyrin and tryptophan at 1518 cm^{-1}, 1552 cm^{-1} and at 1612 cm^{-1} [92]. The remaining 16 clusters had their centroid spectra dominated by a strong fluorescence emission (higher than 10,000 counts), swamping any Raman bands. For the two regions with connective tissue, the k-means cluster analysis provided a good discrimination between the connective tissue and the regions containing larger numbers of viable tumour cells. Similarly, as observed for SCC tissues, the spectra from stroma had a higher fluorescence background than the region with tumour cells, while the Raman bands assigned to nucleic acids had a higher intensity in the regions with a larger number of tumour cells.

In general, spectral differences between the structures in the heterogeneous tissue presented in Figure 3.12 **B**, **C**, **D** and **E** showed that the band from protein (ν(C-C)) and collagen reaching its maximum at 942 cm^{-1} for non-malignant and malignant tissues had shifted its maximum at 958 cm^{-1} and 956 cm^{-1} for the late necrosis region and the mixed region, respectively. We have also compared the contribution of the bands at 1552 cm^{-1} and 1518 cm^{-1} corresponding to porphyrin vibrations for early necrosis, connective tissue and late-stage necrosis. These bands had the highest intensity for the most advanced stage of necrosis, an observation reported in a study that shows that higher contents of porphyrins were observed in necrotic regions of tissues [71].

Comparison normal vs tumour tissues

In Figure 3.13, we display the average Raman spectra for non-malignant lung tissues, adenocarcinoma and squamous cell carcinoma tissue after the subtraction of the fluorescence background. The shaded regions, representative of the standard deviation from the mean, express the intensity variations between the set of values and the mean. The largest intensity variations were observed for non-malignant lung tissue. Subtler variations spread over the entire spectral range were observed for SCC and adenocarcinoma. The spectral variance can be accounted on inter-patient variability and on variations in fluorescence emission, which led to differences in baseline subtraction.

We observed similarities between the average spectra of squamous cell carcinoma and adenocarcinoma with differences in the relative intensity of the bands at 906 - 996 cm^{-1}, 1031 - 1110 cm^{-1}, the band at 1445 cm^{-1} assigned to phospholipids and proteins, and we noticed two peaks specific to squamous cell carcinoma at 1394 cm^{-1} and 1413 cm^{-1}. Besides, the intensity of the bands at 790 cm^{-1}, 814 cm^{-1} and 855 cm^{-1} were higher for non-malignant tissues compared to tumour tissues, as reported in previous studies [35, 36]. A comparable behaviour was observed for the non-malignant samples which presented a strongest signal for the peaks at 1248 cm^{-1}, 1269 cm^{-1} from amide III, collagen, tryptophan and phospholipids vibrations and at 1618 cm^{-1} from tryptophan and porphyrin.

Because the polynomial curve fitting method applied to remove the fluorescence background can lead to spectral artefacts, notably in the case of high level of fluorescence reached [69, 95], we discarded, from now on, the lowest value of signal to noise ratio (SNR <2) for single spectrum. Thus, we obtained a reduced spectral dataset including 2497 spectra from non-malignant tissues, 2518 spectra from adenocarcinoma and 2879 spectra from squamous cell carcinoma.
We used principal component analysis (PCA) as a method to visually represent the spectral differences between normal and malignant lung tissues. Figure 3.14 present the results of the analysis, with the scores and loadings of the two first principal components PC-1 and PC-2 which accounted for 24 % and 7 % of the

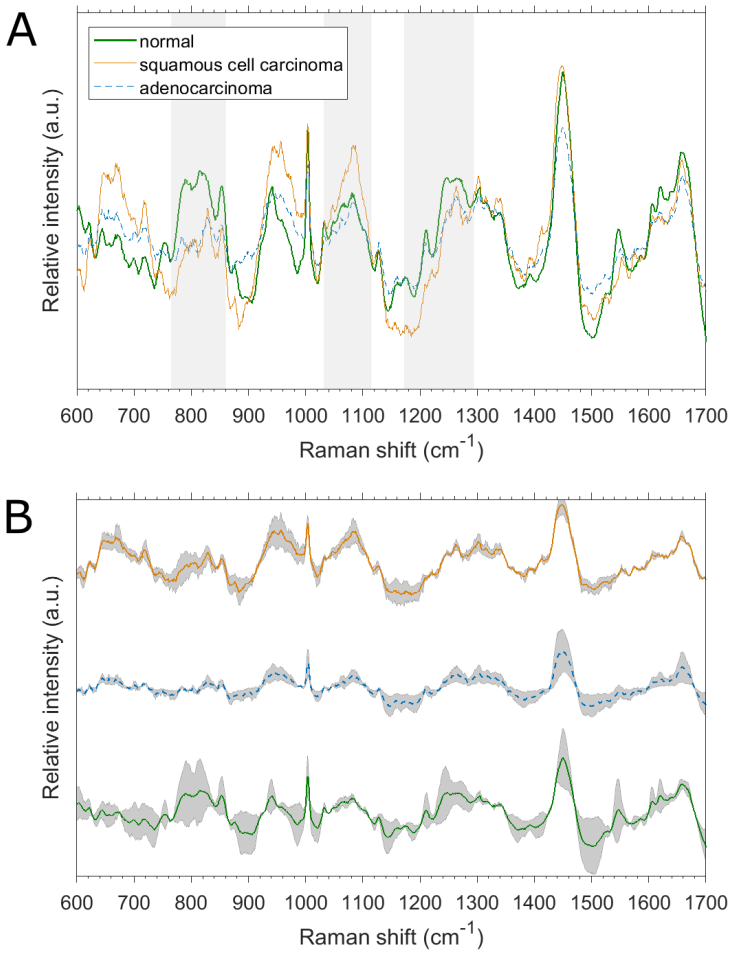
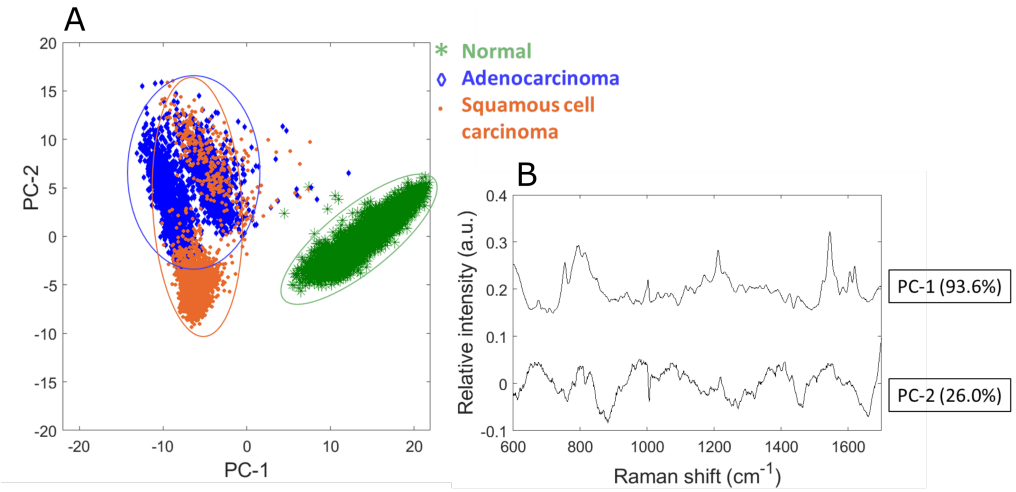


Figure 3.13: **A** Mean Raman spectra of non-malignant lung tissue (green), adenocarcinoma (dotted blue) and squamous cell carcinoma (orange). Each spectrum is the average of around 18,000 spectra from 3 patients (normal), 19,000 spectra from 3 patients (adenocarcinoma) and 16,000 spectra from 5 patients (squamous cell carcinoma). The regions shaded in light grey highlight the main spectral differences between the three types. **B** The shaded regions indicate ±2 standard deviation from the mean. The spectra are represented on the same intensity scale and are offset for clarity.

variance. It is shown that discrimination between healthy tissue, adenocarcinoma and squamous cell carcinoma can be achieved since the PC1 vs PC2 plot shows two distinct clouds of data points that separate non-malignant and malignant tissues. The loadings of the principal components responsible for the spectral separation between the three types of lung tissue contained complex patterns of positive and negative bands, suggesting that the separation is based on a complex molecular pattern. The non-malignant tissues are characterised by positive values of PC-1, including Raman bands assigned to amino acids (1172 cm^{-1}, 1210 cm^{-1} and 1545 cm^{-1}), phenylalanine (1002 cm^{-1} and 1605 cm^{-1}) and haemoglobin

(1620 cm^{-1}). The groups of data from squamous cell carcinoma and adenocarcinoma are situated in the negative part of PC-1 illustrating the contribution from the peaks from proteins (963 cm^{-1}), phospholipids (1092 cm^{-1}) and collagen (1488 cm^{-1}) in their spectra.



C

	Wavenumber (cm^{-1})	Sign	Assignment				
PC-1	815	+	Quartz	**PC-2**	**809**	**+**	**Quartz**
	887	-			829	+	
	963	-	Proteins		883	-	
	1002	+	Phenylalanine		973	+	Collagen, protein
	1092	-	Phospholipids		1006	-	
	1172	+	Tyrosine		1077	+	Phospholipids, DNA and RNA
	1210	+	Tyrosine, phenylalanine		1210	-	Tyrosine, phenylalanine
	1412	-			1220	+	
	1438	-			1270	-	Amide III, collagen, tryptophan
	1488	-	Collagen		1322	-	Collagen and nucleic acids
	1545	+	Tryptophan		1365	+	
	1605	+	Phenylalanine		1390	+	
	1620	+	Porphyrins		1412	+	
					1465	-	
					1552	+	Tryptophan and porphyrin
					1660	-	Amide I

Figure 3.14: **A** Principal component analysis PC-2 vs. PC-1 applied to Raman spectra of non-malignant tissue, adenocarcinoma and squamous cell carcinoma tissues. PC-2 vs. PC-1 shows a discrimination between the spectra from non-malignant samples, sample with adenocarcinoma and sample with squamous cell carcinoma. **B** PC-1 and PC-2, loading plots with corresponding eigenvalues and marker highlighting specific peaks. **C** Assignments of the positive (+) and negative (-) peaks appearing in the loading of PC-1 and PC-2.

To highlight the differences between the squamous cell carcinoma, adenocarcinoma and normal tissues, we also calculated the subtracted spectra by subtracting

the average spectrum of both tumour types by the average spectra from normal tissues, as presented in Figure 3.15.

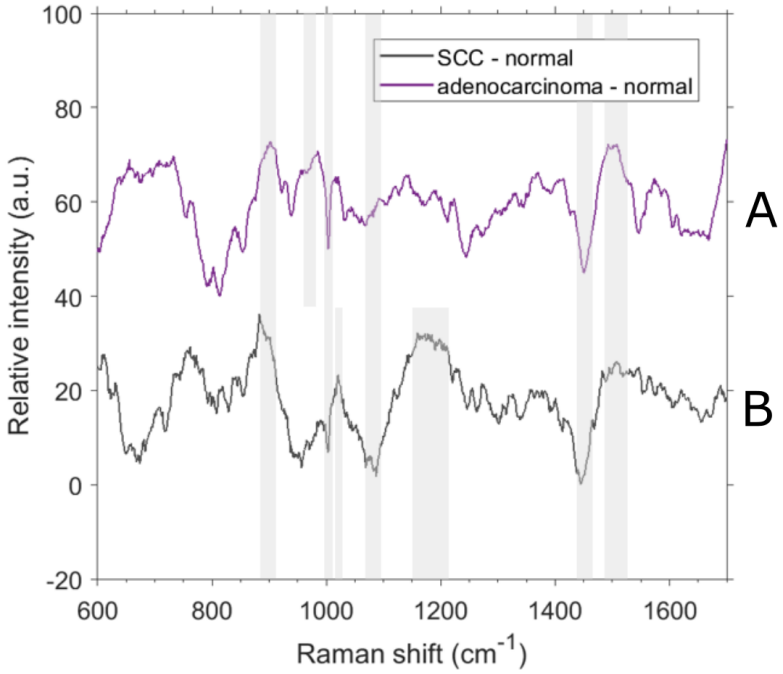


Figure 3.15: Difference Raman spectrum from squamous cell carcinoma minus normal tissue in **A** and adenocarcinoma minus normal tissue in **B**.

The difference spectrum in **A** revealed Raman peaks at 902 cm^{-1}, at 985 cm^{-1}, at 1003 cm^{-1} assigned to phenylalanine, at 1223 cm^{-1} assigned to $\nu(PO_2^-)$, nucleic acids, at 1450 cm^{-1} assigned to collagen and phospholipids and at 1503 cm^{-1}. The difference spectrum in **B** highlights peaks at 883 cm^{-1}, at 1020 cm^{-1}, at 1172 cm^{-1} from tyrosine, at 1450 cm^{-1} assigned to collagen and phospholipids and 1503 cm^{-1} - 1546 cm^{-1} from tryptophan and porphyrin [36, 92].

Discussion

The Raman mapping measurements obtained in this study showed molecular differences between non-malignant, squamous cell carcinoma and adenocarcinoma samples revealed in their Raman spectra. The results presented are consistent with the literature with additional bands observed, not reported before. The differences in the region 1500 cm^{-1} to 1650 cm^{-1} between tumour and normal tissues were explained by stronger bands from porphyrin from blood [96]. It was

reported that higher concentration of porphyrin in tumour tissues is due to lower pH value around the tumour [97]. The porphyrin, in a different local environment, could also experience protonation. In fact, some lung cancer detection methods are based on the detection of levels of porphyrin and tryptophan in body fluid indicative of cancer; the higher the level the most advanced stage of cancer [98]. Furthermore, necrotic tissue can be discriminated from healthy tissue and region with living tumour cells by the intensity of porphyrin peaks at 1518 cm^{-1} and 1552 cm^{-1} and ratio of intensity at 856 cm^{-1} and 829 cm^{-1}. It confirms the finding of Croce et al. which was explained by tumour-specific metabolism with porphyrin synthesis failure [99, 100].

The bands at 1450 cm^{-1} and at 1666 cm^{-1} change appearance for malignant tissue, an observation also reported by Kaminaka et al [35].

The difference spectra between tumour and normal tissues presented in Figure 3.15 for squamous cell carcinoma was in good agreement for bands at 1078 cm^{-1}, at 1450 cm^{-1} and at 1552 - 1582 cm^{-1} with the conclusions from Huang et al. It confirms that squamous cell carcinoma is associated with higher signals from collagen, tryptophan and phospholipids.

In addition, most spectra shared a common feature: the collection of an intense fluorescence background. We expect that most fluorescence background from lung tissues measured comes from mixed sources from which we can count haemoglobin, elastin fibres, collagen, tryptophan and porphyrins. We noticed in this study that a higher proportion of healthy lung tissues were involved in the observation of an extremely high level of fluorescence detection. Visual examination of most of these tissues (appearing more saturated in red colour under light microscopy) often seemed to exhibit higher level of blood than tumour tissues matching the assumption from Tang et al [101]. This observation could be explained by fluorophore environment variations in normal cells compared to cancer cells. Incident laser light could be absorbed more by blood which could lead to an increase emission of signal from haemoglobin. This emission would be interfering with Raman signal leading to spectra overshadowing or fluorescence detection saturating the detector. On the other hand, elastin is a protein found in lung tissues (alveolar walls, arteries) for tissue elasticity [102, 103] that has a natural autofluorescence at 488 nm. This may be related to other processes such

as ageing and cancer development which could disrupt the elastin composition explaining the patient variability in terms of fluorescence levels.

To envision the role of Raman spectroscopy for clinical applications, the tissue metabolism should be understood to correct the translation of the results from *ex vivo* to *in vivo*. We were aware that after resection and Raman measurement, the changes in metabolism may induce differences in results. We will discuss in the following section ways to address the issues encountered during the study.

Perspectives

This study on lung tissue has been mostly limited by the complexity of lung tissue behaviour under measurement conditions and histopathological staining. As we noticed, few tissues samples were sensitive to laser irradiation and room temperatures leading to tissue damage with no explanation. Also, H&E sections could appear distorted compared to their adjacent section making the correlation between the two sections challenging and forcing us to discard the spectral data obtained from these tissues. There is a need in developing a method which would overcome the poor structural information obtained from the bright-field images. Actually, we have observed the lack of contrast, necessary to distinguish features, of the microscopic images collected when a lung sample is slightly thicker than usual. For this reason, we have investigated the use of autofluorescence imaging that could help in tissue microstructure detection. The imaging system, a Nikon C2 confocal microscope, was connected to the second port of the currently used microscope. The laser unit OBIS with an excitation wavelength of 405 nm was focused through a Plan-APO 2x objective and the fluorescent light was collected through a 511 nm longpass filter (FF511-Di01, Chroma).
Figure 3.16 illustrate how the autofluorescence image reveals internal tissue features. Study on this tissue would have been compromised if we only had access to the bright-field image seen in (c) showing only the tissue outline that differs significantly from the outline of its adjacent H&E stained section in (a). The uncertainty may be significantly reduced with the autofluorescence image

that reveals structures and cellular morphology such as normal alveolar tissue structures and blood vessels making it a straightforward approach to correlate the tissue with its H&E [104].

It could be envisioned to use the autofluorescence imaging technique to understand and measure tissue alteration after Raman measurements. The comparison between images in **C** and **D** provides an indication of the tissue transformation induced by the two Raman scanning measurements performed, showing tissue structural modification, thinning of alveolar spaces, and modification of the vessel's appearance explaining tissue shrinking.

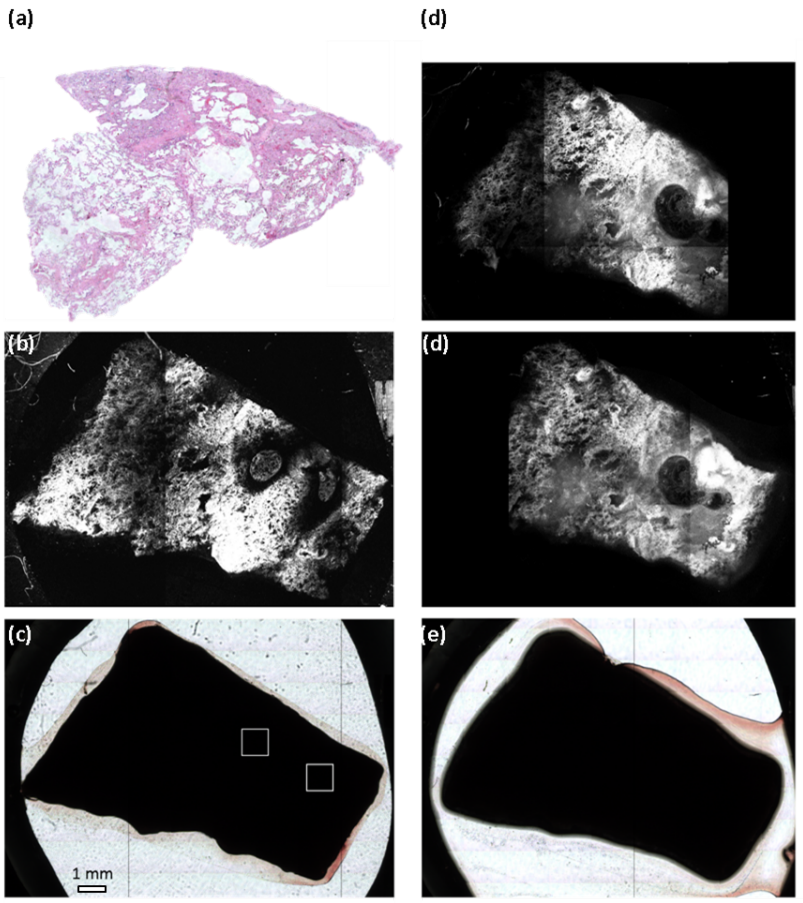


Figure 3.16: **A** Adjacent lung section stained with Haemotoxylin and Eosin of the tissue shown in **C**. **B** Autofluorescence (AF) image of the lung tissue before Raman measurement. Two Raman maps were acquired in the region displayed in **C** for a total of two hours and thirty minutes. The microscopic image of the tissue, after the Raman measurements, is shown in **E** with AF images in **D** at two different focus depth.

In summary, 18 resected human lung tissue samples were investigated by Raman mapping measurements at 785 nm. 60 % of the samples presented well-defined Raman bands that were correlated with specific tissue structures from the H&E assessment. The samples consisted of non-malignant tissues as well as tissues with squamous cell carcinoma and adenocarcinoma. Tissue types were fully characterised by the variation in intensity of the Raman bands from nucleic acids, proteins, amino acids, phospholipids, phenylalanine, tryptophan, and haemoglobin and the intensity of fluorescence. We even reported spectral differences for tissue specificities such as the stage of tissue necrosis. We also noticed tissue shrinkage and fluorescence level variability between patients. The first observation was not correlated to any specific feature of the tissue, while the second seemed to be correlated with a higher level of blood of the tissue. Finally, and for most tissues, the Raman signal was collected with an intense fluorescence background.

The following chapter presents modifications of the original Raman spectrometer design to help us overcome the separation between the Raman signal from the fluorescence background.

4

Characterisation of human tissue samples by polarised Raman spectroscopy

In this chapter, polarised Raman spectroscopy mapping was used to obtain spectral images of human tissues. We investigated the difference between Raman spectra of normal dermis in various morphological configurations. We also studied the effect of varying the polarisation angle for skin and two tissue types exhibiting a high level of laser-induced fluorescence emission.

4.1 Introduction

Studies in the literature on skin tissue research indicate that Raman spectroscopy was used as an imaging tool of cells [105], for tissue diagnosis [106] and for drug mechanism understanding [107]. Diagnosis applications include the investigation of different types of skin neoplasms mainly cancers of keratinocytic origin. Early work showed that basal cell carcinoma can be differentiated from the healthy surrounding tissue *in vitro* by Raman spectroscopy mapping [80]. Raman mapping allowed the authors to detect a lower collagen content and lower amount of lipids and nucleic acids in the dermis surrounding the tumour. They also observed that BCC and epidermis had similar Raman spectra making them hard to discriminate. Lieber et al. were able to classify BCC, squamous cell carcinoma and melanoma from 39 patients at different measurement depths [108]. Larraona-Puy et al. showed, based on an automated discrimination

model, the feasibility of discriminating BCC from dermis and epidermis with 90 % sensitivity and 85 % specificity [109]. The sensitivity of a classification model is the proportion of true cases correctly identified, while the specificity is the proportion of non-cases correctly identified. Recently, the classification model developed by Kong et al. allowed the detection of BCC in 20 unseen samples with 95 % sensitivity and 94 % specificity [110]. The feasibility of such successful studies is encouraging but false positive classifications were found emphasising the difficulty in discriminating BCC, epidermis, hair follicles and inflammation [109]. Besides, studies also reported the collection of a superposition of the Raman peaks and fluorescence background of the skin tissue limiting the spectral analysis information [111]. To address the current issues, we proposed to use polarised Raman mapping spectroscopy. First of all, for a better understanding of the tissue chemical composition at a microscopic level. Taking into consideration that most molecules scatter light in both parallel and perpendicular planes since their polarizability is anisotropic [112], the emitted Raman scattered radiation is likely to be in a different plane than the incident polarisation. Instead of using an incident polarised laser light to excite the tissue sample and to collect the emitted Raman scattered light as in classical Raman experiments, we will vary the direction of the polarised light. By doing so, the variation of the plane of the emission and the Raman light for a certain vibration in the fingerprint region will be quantified using the depolarisation ratio ρ, defined as the ratio in intensity of light in the perpendicular direction to the one in parallel. This quantity expresses the polarisation characteristic of the Raman emission and is useful for a deeper understanding of the tissue molecular orientation and symmetry of bands. Few studies have reported using polarised Raman spectroscopy for human skin tissue investigation. Ly et al. showed a structural alteration of collagen in peritumoral stroma [60, 113]. However, to our knowledge, no Raman mapping polarised measurements were ever reported. For this reason, we investigated spectral variation occurring within the polarised Raman maps of skin tissue. We aim, with the maps obtained, to correlate with precision the map with its corresponding H&E stained image. It will allow us to select spectra with morphological specificities within the same map for the study of molecular variations at the microscopic level. Furthermore, the molecules responsible for the fluorescence emission in a

biological tissue might present preferential orientation within the tissue leading to Raman light in specific directions. This hypothesis will be tested for the case of fluorescent tissues where the fluorescence background is difficult to separate from the Raman signal by classical Raman spectroscopy.

4.2 Material and Methods

Sample preparation

Skin tissue samples were obtained during Mohs' micrographic surgery at the Nottingham NHS Treatment Centre. Excised tissues were frozen and cut in two adjacent sections. The 10 μm thick first section was stained by haematoxylin and eosin (H&E) and used as the standard of reference for the diagnostics. The remaining tissue section was defrosted, washed and kept frozen at – 20 °C. The preparation of the tissue for Raman mapping measurement involved to immerse it in PBS and leave it to defrost at room temperature for a few minutes. This process allowed the tissue to reach the ambient temperature and to undergo a final washing to avoid any remaining blood. The skin tissue was then placed on the 100 μm thick quartz substrate. The tissue is immersed in PBS with its surface covered with cotton to avoid it drying. When the tissue was ready for investigation, it was securely placed on the microscope slide holder. We first used a 4x objective (Olympus, Plan N 0.10) to obtain a full image of the tissue under bright-field conditions. The bright-field image reveals the tissue contours and highlights the blue and red ink staining useful to us for tissue orientation. In some cases, the image showed residual staining artefacts that would lead to the emission of sharp Raman peaks and high fluorescence hiding the actual Raman information of interest. The image also allowed the identification of some tissue features such as sebaceous gland and the contrast, brightness and texture differences are useful to correlate with the H&E image given by the histopathologist. Once the area of interest was selected, we proceeded to the Raman spectroscopic measurements.

Instrumentation

The polarised Raman experiments were carried out using a lab-built near infrared polarised Raman micro-spectrometer described in chapter 2 section 2.2. The parallel-polarised spectra were obtained when $\theta = 0^\circ$ meaning the polarised laser light is oriented parallel to the direction of the analyser. In contrast, $\theta = 90^\circ$ expresses the perpendicular configuration. After passing through the analyser, the light was focused via an achromatic lens to the entrance of a 50 μm of diameter optical fibre connected to the spectrograph. The focus was adjusted to maximize the collection of in-focus Raman scattered light.

Data acquisition

We performed polarised Raman mapping measurements of different sizes with a maximum width of w = 2 mm and length l = 1 mm. The spectra were recorded at an increment of 20 μm and were measured at each point for 2 s. The laser power at the sample was equal to around 70 mW in all experiments which will not induce any skin damage. The maps were measured at various collection angles θ, even though the laser is not leading to tissue damage, we preferred to randomize the order at which the measurement was taken to avoid any effect of this order. The data we present was acquired for each tissue within a 6-hour period after tissue being thawed.

Data processing

The depolarisation ratio of a band i, $\rho_{(i)} = I_{\perp(i)}/I_{\parallel(i)}$, was calculated as the ratio of the intensity of the band area in the cross-polarised configuration by the intensity of the band area in the parallel-polarised configuration. The band area was calculated by selecting two points (A,B) on both sides of the peak from which a line will be fitted. This is to correct the influence of the fluorescence background present in the signal. Two other points (C,D) corresponding to the actual peak of interest had their baseline, calculated previously, subtracted to

provide an accurate area under band. The method is illustrated for two bands in Figure 4.1.

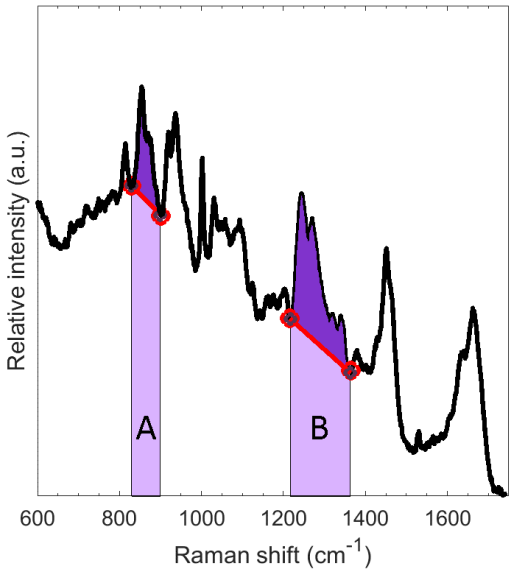


Figure 4.1: Parallel-polarised spectra of healthy skin dermis. The light purple shaded area indicates the region for the selection of the baseline fit and the darker purple highlights the spectral region that are considered for the area under the peak between 829.2 cm^{-1} - 899.9 cm^{-1} in **A** and 1217 cm^{-1} - 1363 cm^{-1} in **B**.

The selection of the points has been optimised and was consistent throughout all the study for all the collection angles. We also present the principal component analysis (PCA) based images. They are based on k-means clustering, a decomposition process which group together similar structural features in the data set. It allows a simple and standard distinction of different spectra classes of a biological sample each represented with a colour.

4.3 Results

In this part of the experiment, a total of fifteen skin samples from seven different patients were analysed using polarised Raman spectroscopy.

Sample ID	Tissue type	Date of measurement
Sample A	BCC, inflammation	10/11/17
Sample B	Inflammation, inflammed dermis, BCC	15/11/17
Sample C	BCC, dermis	21/11/17
Sample D	Dermis	22/11/17
Sample E	BCC, dermis	19/01/18
Sample F	BCC, inflammation, dermis	18/01/18
Sample G	BCC, inflammation, sebaceous gland	16/01/18
Sample H	BCC, inflammation, sebaceous gland, dermis	22/01/18
Sample I	Dermis, inflammation	05/02/18
Sample J	Dermis, inflammation, BCC	16/11/17
Sample K	Dermis, sebaceous gland, epidermis	17/11/17
Sample L	Dermis	01/06/18
Sample M	Dermis, epidermis, sebaceous gland	08/06/18
Sample N	Dermis, BCC, sebaceous gland	04/06/18
Sample O	BCC	22/06/18

Figure 4.2: Table summarizing the skin tissue samples considered in the study. Adjacent lines with the same colour background indicate tissue sample from the same patient but from different regions. BCC stands for basal cell carcinoma.

Characterisation of skin dermis

Figure 4.3 shows the result of Raman parallel-polarised and cross-polarised spectra of healthy skin connective tissue from two patients. The spectral contribution of the quartz substrate was subtracted from each spectrum. No other processing was applied and no offset was applied to the y-axis. The spectra in parallel configuration showed the notable peaks reported in literature [92]. Peaks at 852 cm^{-1}: CCH bending of protein, 939 cm^{-1}: C-C stretching of protein (backbone), 1004 cm^{-1}: C-C vibration of phenyl ring, 1246 cm^{-1}, 1254 cm^{-1}, 1269 cm^{-1}: amide III, C-N stretching and N-H in plane bending, 1298 cm^{-1}: C-H modes of lipids, 1336 cm^{-1}: elastin, 1337 cm^{-1}: DNA & RNA base, 1378 cm^{-1}: linear stretch of C-C bonds, 1450 cm^{-1}: C-H bending of proteins, C-H stretching and assymetric,1653 cm^{-1}: C-O stretch of amide I, 1656 cm^{-1}: C-C lipids and 1665 cm^{-1}: C-O amide

I vibration were observed. In perpendicular, we observed a drastic reduction of peaks intensity below 1004 cm^{-1} and, in the spectral interval 1000 - 1700 cm^{-1}, a flattening of the fluorescent background. Spectra at both angles of collection clearly illustrate a strong polarisation dependence of Raman bands.

We were also interested in measuring several tissue regions for each sample to account for inter-patient variability. As observed in Figure 4.3 **E** compared to Figure 4.3 **F**, a large patient variability was noted in terms of fluorescence level. Indeed, spectra collected at less than 0.5 mm away from another position presented a stronger fluorescence background with weaker Raman peaks. We also can highlight from the two H&E images the difference in size of the epidermis region. The skin specimens obtained during Mohs' surgery are sectioned in a way such that the thickness of the tissue is larger on the side where the ink will be applied and the opposite edge the section is thinner. On top of this consideration, the epidermis layer is in general 20 to 60 times thinner than the dermis layer for the types of skin most commonly involved in basal cell carcinoma occurrence. These considerations could justify the cotton bands appearing at 746 cm^{-1}, 1337 cm^{-1} and 1528 cm^{-1}.

In this study we were particularly interested in the molecular properties of dermis, as previous works [80, 114] indicated changes in collagen structure around BCC. In the following part of the study we investigated the spectral characteristics of healthy dermis depending on its localisation within the tissue.

We present in Figure 4.4, polarised Raman spectra of healthy dermis nearby basal cell carcinoma. The spectra from the BCC sample, in parallel configuration, revealed characteristic peaks of BCC are observed at 1085 cm^{-1} from lipids and nucleic acids backbone (PO_2 symmetric stretching), 1127 cm^{-1} from C-C stretching of lipids, 1207 cm^{-1} from tyrosine and phenylalanine, 1305 cm^{-1} from lipids and amide III, 1336 cm^{-1} from adenine, phenylalanine and CH deformations consistent with the literature [80]. For $\theta = 90°$, for dermis close to BCC and BCC itself, we noted the strong reduction of the peaks below 1004 cm^{-1}. We also

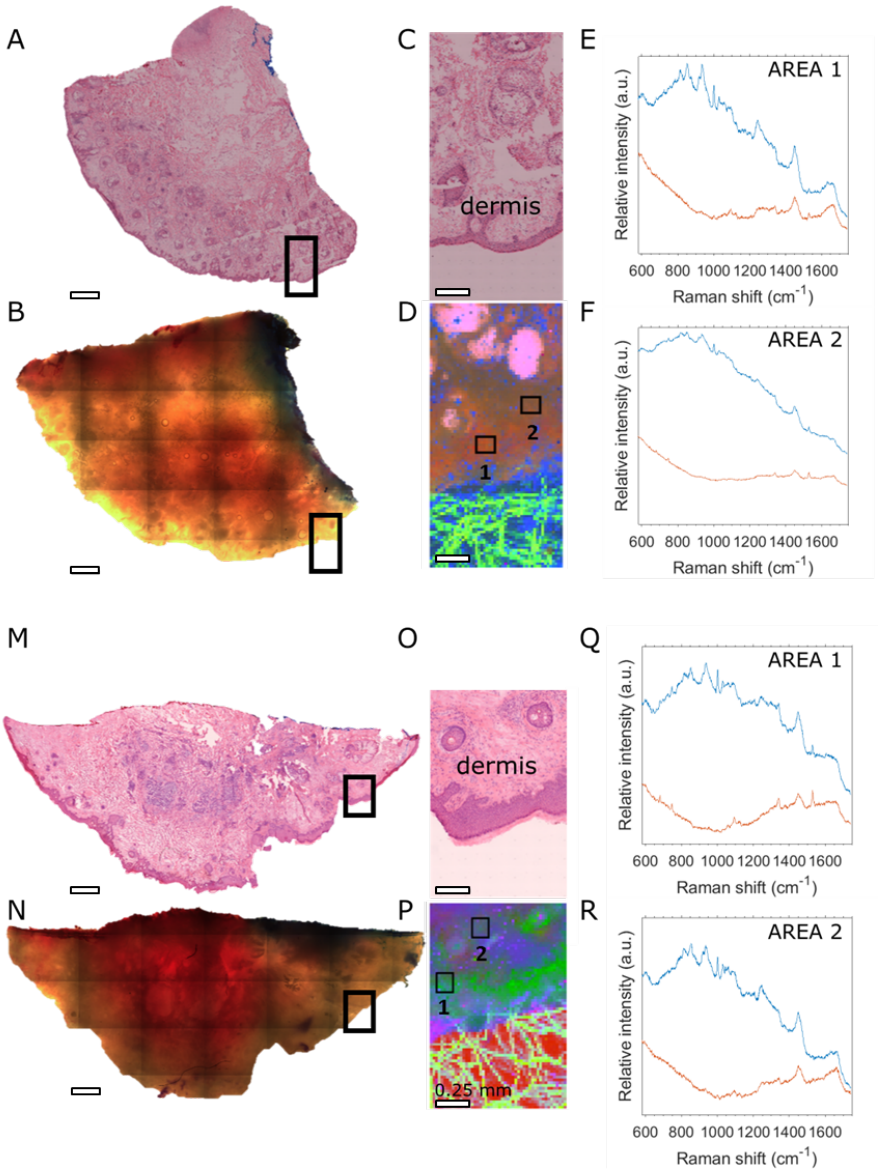


Figure 4.3: Investigation of two skin samples presenting healthy dermis by polarised Raman spectroscopy. We recorded microscopic bright-field images of the entire samples in **B** and **N**. The corresponding H&E images **A** and **M** provided a detailed diagnosis of the tissue and allowed us to select area of interest as represented by the black rectangles. In **C** and **O** are the zoomed-in views of the area. We then proceeded to the 1 mm x 2 mm and 1 mm x 1.5 mm Raman mapping measurements. **D** and **P** display the PCA images obtained. We chose within the maps two regions of around 60 spectra represented by two numbered rectangles in the PCA image. We show in **E**, **F** and **Q**, **R** the average spectra from the two regions mentioned in blue for the parallel configuration and in red for the perpendicular configuration. Scale bars indicate 1 mm in **A**, **B**, **M** and **N** and 0.25 mm everywhere else.

noticed that the amide I band at around 1660 cm^{-1} had equal or even higher intensity count for BCC and dermis in perpendicular compared to parallel.

Concerning dermis in the vicinity of BCC, the Raman bands were weak compared to the fluorescence level making it challenging to report specificities of the spectra. We observed a higher level of fluorescence background for the sample in Figure 4.4 **M**. This sample was comprised of nests of tumour cells in the dermis while the sample in Figure 4.4 **A** contained islands of tumour cells in a dense fibrotic stroma.

Next, we investigated the spectral variations in dermis close to a region with inflamed tissue. Figure 4.5 shows the study of polarised Raman measurement of inflammation and healthy dermis close to an inflamed region.

In the parallel-polarised spectra from the dense inflamed region, similar spectral variations are discerned, comparable to the ones for the healthy dermis. However, we distinguished some particularities in the wavenumber range 1200 - 1400 cm^{-1}. Both samples represent inflammation in the superficial dermis and the spectral differences suggested a variation in the nucleic acid composition and protein content. The perpendicular-polarised spectra from the inflamed region were similar than reported before for BCC and healthy dermis in the lower wavenumber region and, for the first region of the tissue **A**, the 1298 cm^{-1} assigned to lipid vibration is strongly pronounced [92].
For dermis in the vicinity of inflammation, the perpendicular configuration did not seem to reveal any information. Once again, the spectra in parallel highlighted the patient variability in terms of fluorescence intensity emission. We observed for the parallel spectra presented in **Q** a higher intensity for the peak at 1125 cm^{-1} attributed to vibration of acyl backbone in lipids.

Overall, this part of the study emphasises that Raman spectra from skin tissue had a fluorescence background which varies from patient to patient and within a same tissue. Also, we noticed that the spectra collected from the tissues in perpendicular did not carry as much spectral information as the ones collected in the parallel configuration. For that reason and from now on, perpendicular spectra will not be analysed on their own but will be combined to the spectra in

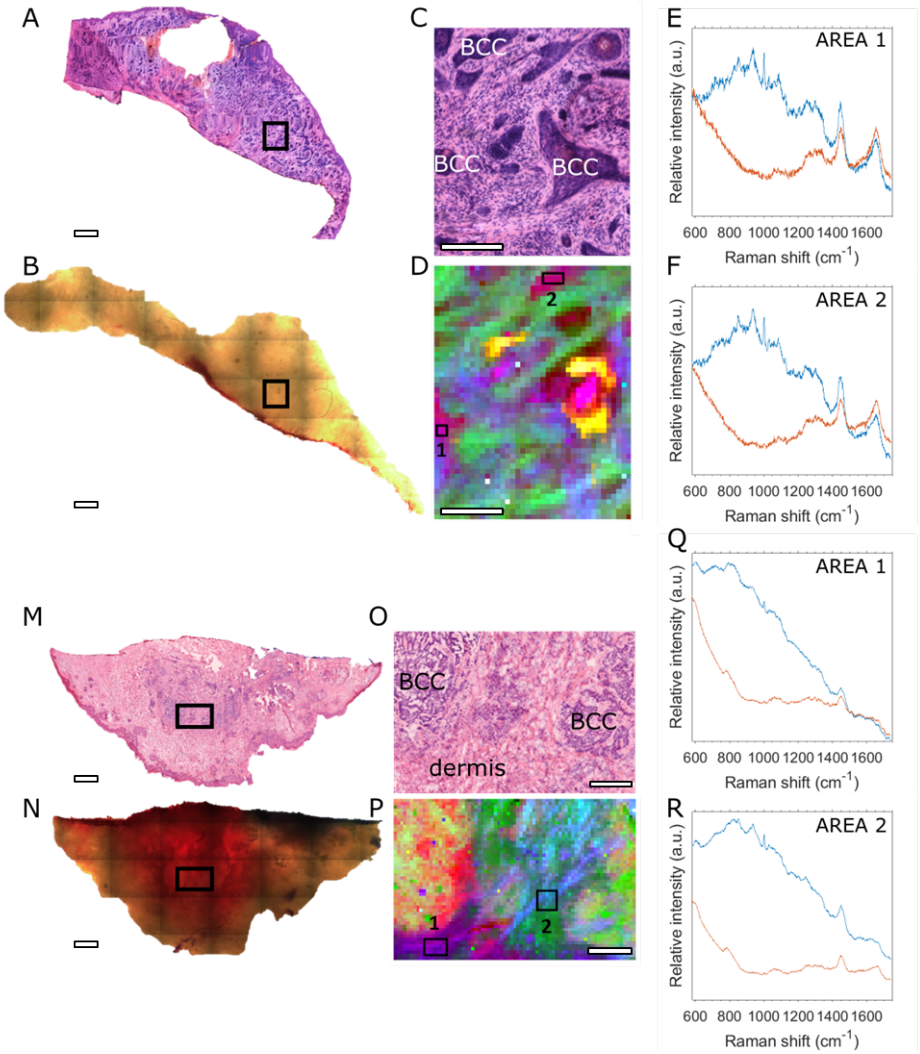


Figure 4.4: Polarised Raman spectroscopy of two samples presenting BCC and dermis in the vicinity of BCC. We recorded microscopic bright-field images of the entire samples in **B** and **N**. The corresponding H&E images **A** and **M** provided a detailed diagnosis of the tissue and allowed us to select areas of interest as represented by the black rectangles. In **C** and **O** are the zoomed-in views of the area with highlighted the region where BCC and dermis are found. We then proceeded to the 0.8 mm x 1 mm and 1.5 mm x 1 mm Raman mapping measurements. **D** and **P** display the PCA images obtained. We chose within the maps two regions of around 10 spectra for the top sample and 50 spectra for the bottom one represented by two numbered rectangles in the PCA images. We show in **E**, **F** and **Q**, **R** the average spectra from the two regions mentioned in blue for the parallel configuration and in red for the perpendicular configuration.

parallel configuration.

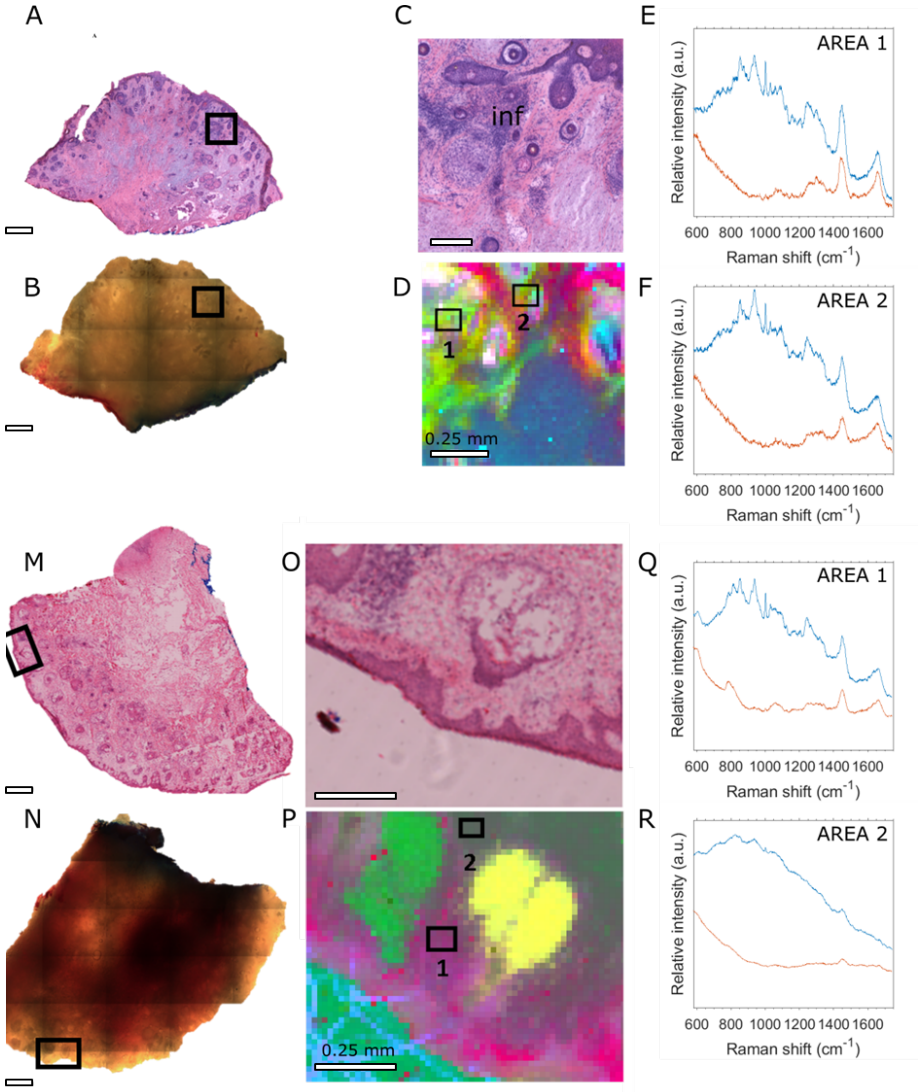


Figure 4.5: Polarised Raman spectroscopy of two samples presenting BCC and dermis in the vicinity of BCC. We recorded microscopic bright-field images of the entire samples in **B** and **N**. The corresponding H&E images **A** and **M** provided a detailed diagnosis of the tissue and allowed us to select area of interest as represented by the black rectangles. In **C** and **O** are the zoomed-in views of the area. We then proceeded to the 1 mm x 1 mm and 1.5 mm x 1 mm Raman mapping measurements. **D** and **P** display the PCA images obtained. We chose within the maps two regions of around 40 spectra for the top sample and 30 spectra for the bottom one represented by two numbered rectangles in the PCA images. We show in **E**, **F** and **Q**, **R** the average spectra from the two regions mentioned in blue for the parallel configuration and in red for the perpendicular configuration.

Then, the spectral variations of skin dermis in different morphological configurations were quantified with the calculation of the depolarisation ratio.

Depolarisation ratio

To quantify the modification of the ordering and structure of the molecules occurring in the dermis caused by cancer or inflammation, we calculated the depolarisation ratio of five specific bands corresponding to various chemical vibrations. Depolarisation ratios were calculated from the integrated area of the peaks of the polarised Raman spectra recorded in the perpendicular and parallel configurations. Depolarisation ratio values range from 0 to 2, in our case, indicating values of highly polarised bands for the lowest values to depolarised bands for the highest ones. Studies in the literature reported collagen degradation and changes in protein secondary structure in dermis surrounding disease regions [60, 114], for this reason we chose few bands assigned to protein vibrations to verify the previous observations.

We selected bands responsible of protein vibrations at 937 cm^{-1} from protein backbone, at 1249 cm^{-1} from amide III and at 1660 cm^{-1} from amide I, as well as bands from lipids at 1450 cm^{-1} from CH_2 vibrations and nucleic and amino acids vibrations at 856 cm^{-1}.

The integrated areas of the bands were calculated between 829 cm^{-1} and 900 cm^{-1} for the band at 856 cm^{-1}, between 900 cm^{-1} and 985 cm^{-1} for the band at 937 cm^{-1}, between 1217 cm^{-1} and 1363 cm^{-1} for the band at 1249 cm^{-1}, between 1413 cm^{-1} and 1504 cm^{-1} for the band at 1450 cm^{-1} and between 1542 cm^{-1} and 1716 cm^{-1} for the band at 1663 cm^{-1}.

In Figure 4.6 are displayed the values of the depolarisation ratio obtained for each tissue type. Each bar of the histogram represents a tissue sample, we included a total of 13 patients for which we calculated the depolarisation ratio value for the bands mentioned earlier. We represent the average values along with the value obtained for each spectrum included in the area considered.

The value variations between samples from the same type can be justified by the morphological specificity of the tissues. As illustrated in Figure 4.7, even though the two samples represent homogeneous healthy dermis from normal skin stroma, the region of dermis measured in the first sample (**A**) was situated near a large area of dense fibrous connective tissue while the region from the second tissue

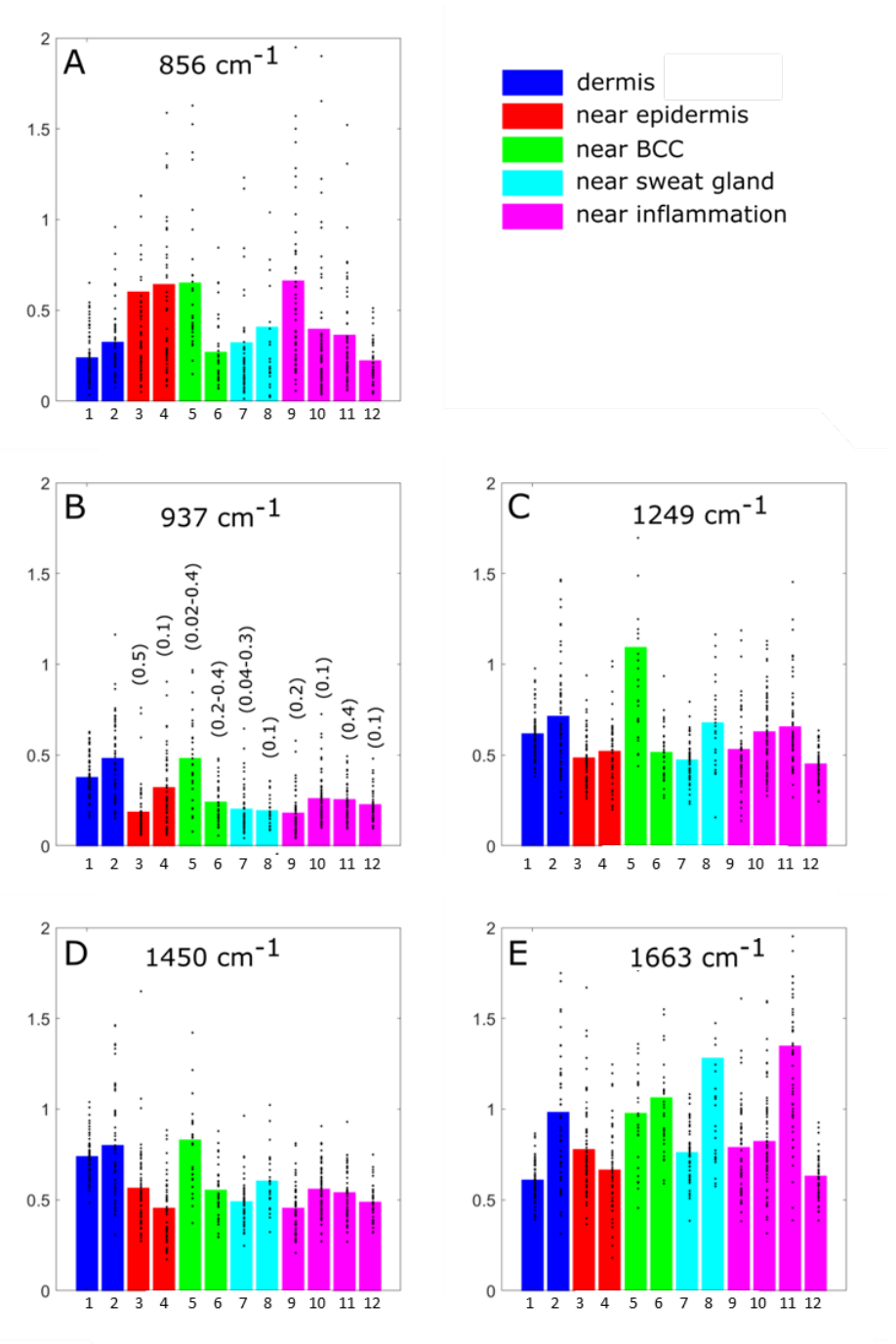


Figure 4.6: Depolarisation ratio of the bands at 856 cm^{-1} (**A**), 937 cm^{-1} (**B**), 1249 cm^{-1} (**C**), 1450 cm^{-1} (**D**) and 1663 cm^{-1} (**E**) for each tissue morphology. Each bar (numbered 1 to 12) is the average value of the depolarisation ratio, each individual data points (in black) are superimposed to the histogram bars. The first two bars represent healthy dermis (dark blue), bar 3 and 4 represent dermis in the vicinity of epidermis (red), bar 5 and bar 6 for region near BCC (green). Bar 7 and 8 for dermis near sebaceous gland (cyan) and the final 4 bars dermis near an inflamed region (magenta). Distances in mm between dermis and the different features are indicated in bracket in **B**.

(**B**) was near the hypodermis: a layer of adipose tissue.

Overall, we observed a slight variation of the average values within each group except for the two BCC tissues. Indeed, for four of the bands presented, the depolarisation ratio fluctuated drastically between the tissues. The H&E of these tissues presented in Figure 4.7 **C** and **D** revealed that the dermis region selected from the second tissue was also in the vicinity of an inflamed region which could explain that the value of the ratio tends to be closer to the group of samples near inflammation.

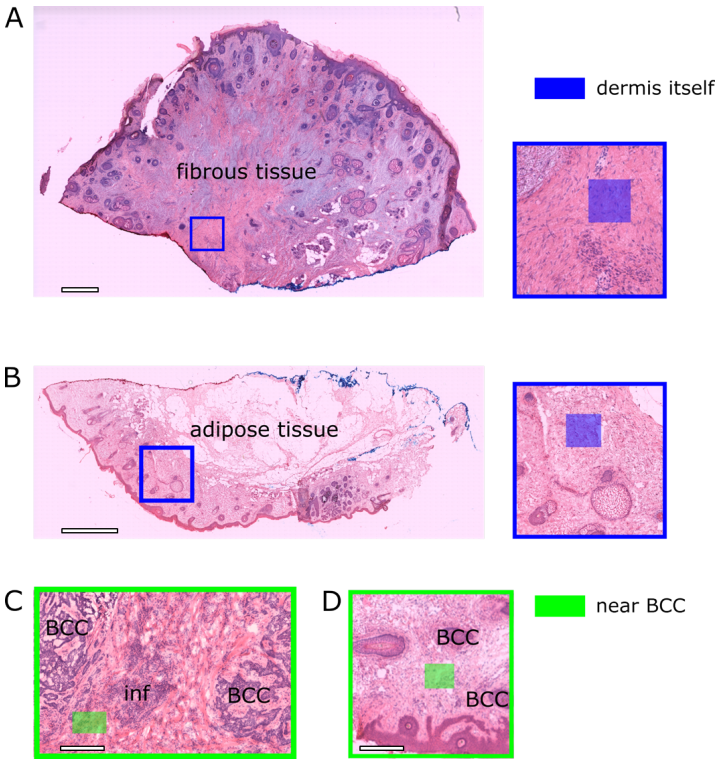


Figure 4.7: Microscopic images represent the H&E stained images of the healthy dermis samples **A** and **B** considered in this experiment. The right panel represent the area measured with the blue rectangle highlighting the region from which the spectra where acquired. Scale bar indicates 1 mm. **C** and **D** display the H&E of the tissue samples accounting for dermis near BCC with the green rectangle illustrating the region of dermis chosen. Inf stands for inflammation. Scale bar indicates 0.25 mm

The depolarisation ratio of the band at 856 cm^{-1} reaches its lowest value for dermis alone. The highest value is attained for epidermis. For the band 937 cm^{-1}, the average depolarization ratio ($\rho_{937(\text{dermis})} = 0.43$) of the two healthy dermis tissues is higher than the average values for dermis close to any morphological features studied. This band is assigned to the collagen backbone vibration so this

result could indicate collagen modification in the dermis in the vicinity of a morphological feature such as sebaceous glands, BCC, inflammation or epidermis. It was reported in previous studies that collagen structural changes occur in dermis close to tumour [114], suggesting an increased content of collagenase [115]. The depolarisation ratio for the band at 1249 cm^{-1} for most groups fluctuated around $\rho_{1249} = 0.5$. For healthy dermis ($\rho_{1249(dermis)} = 0.66$) while for dermis next to BCC ($\rho_{1249(peritumoral)} = 0.80$) which are included in the range of values recorded in the literature [60]. For the 1450 cm^{-1} band, the average depolarisation ratio for dermis alone was higher ($\rho_{1450(dermis)} = 0.77$) than for all the other types. The depolarisation ratio values of the band at 1663 cm^{-1}, for every tissue types were higher than $\rho_{1663} > 0.75$ indicating depolarised vibrations (not totally symmetric vibrations) except for epidermis for which $\rho_{1663} = 0.72$.

In summary, skin tissue comprises extensive and complex molecular architecture which makes the classification of a tissue region into a single category challenging. Variation of the depolarisation ratio values account for these differences which reflect tissue inherent morphological structure variety and the influence of distance from a feature from which the spectra were considered. Nevertheless, we reported that for normal dermis, there is a higher level of alignment (smaller depolarisation ratio) for nucleic and amino acids (band at 856 cm^{-1}) compared to the other groups, especially compared dermis in the vicinity of epidermis for which the depolarisation ratio is the highest for this band. On the other side, the depolarisation ratio of the bands at 937 cm^{-1} and 1450 cm^{-1} clearly showed a difference in the molecular alignment of the collagen backbone and lipids between dermis alone and dermis close to any feature.
For a closer understanding of the polarisation dependence, the polarised Raman study was extended to the investigation of various angles of collection of the polarised Raman spectra.

Dependency of the Raman bands on the polarisation angle

Results in Figure 4.8 **A** and **B** show that some bands experience a large polarisation dependence, for example the bands at 937 cm^{-1} and 1663 cm^{-1}. We also plotted

the calculated area under the bands as a function of the collection angle θ. The area for the band at 937 cm^{-1} showed for both tissue types a minimum of intensity when the analyser is oriented perpendicularly compared to the incident laser light, and a maximum when in parallel configuration. The depolarisation ratio of the band area at 1450 cm^{-1} showed a similar behaviour for both samples, it presents a maximum at the angle $\theta = 0°$ for healthy dermis and BCC, and a minimum at $\theta = 120°$ for both. The depolarisation ratio at 1663 cm^{-1} had comparable values for basal cell carcinoma and dermis. However, for this band, ρ_{1663} had a maximum at $\theta = 40°$ for both tissue types. The results described above brought our attention to the fact that additional information on the nature of the polarizability associated to these molecular vibrations can be obtained at collection angles different from the commonly used $\theta = 0°$ and 90°.

Hence, the study was extended to a full rotation of 360°. Single spectra were recorded with the polariser rotated in steps of 40° for the same two samples. The peak intensity values for BCC and healthy dermis are represented in the polar diagrams in Figure 4.9. The band intensities were normalised to the value in the parallel configuration ($\theta = 0°$).

The shape of the distributions indicated a strong anisotropic scattering especially for the band at 937 cm^{-1} and at 1663 cm^{-1} for both tissue types. The band at 937 cm^{-1} from proline, hydroxyproline and collagen backbone vibrations had a behaviour which reflects a preferential orientation of the molecules in the parallel direction where the value reaches the maxima in the polar diagram. However, for the band at 1663 cm^{-1} from amide I, phenylalanine, tyrosine and tryptophan we observed maxima in the 40° orientation. It implies that the preferential polarizability of the group of molecules responsible to the 1663 cm^{-1} band was turned 40° compared to the one from the group of molecules for the 937 cm^{-1} band. The band at 1450 cm^{-1} showed less preferential orientation.

From the previous observations, we noticed that the 40° angle collection revealed interesting polarisation characteristics never reported in the literature

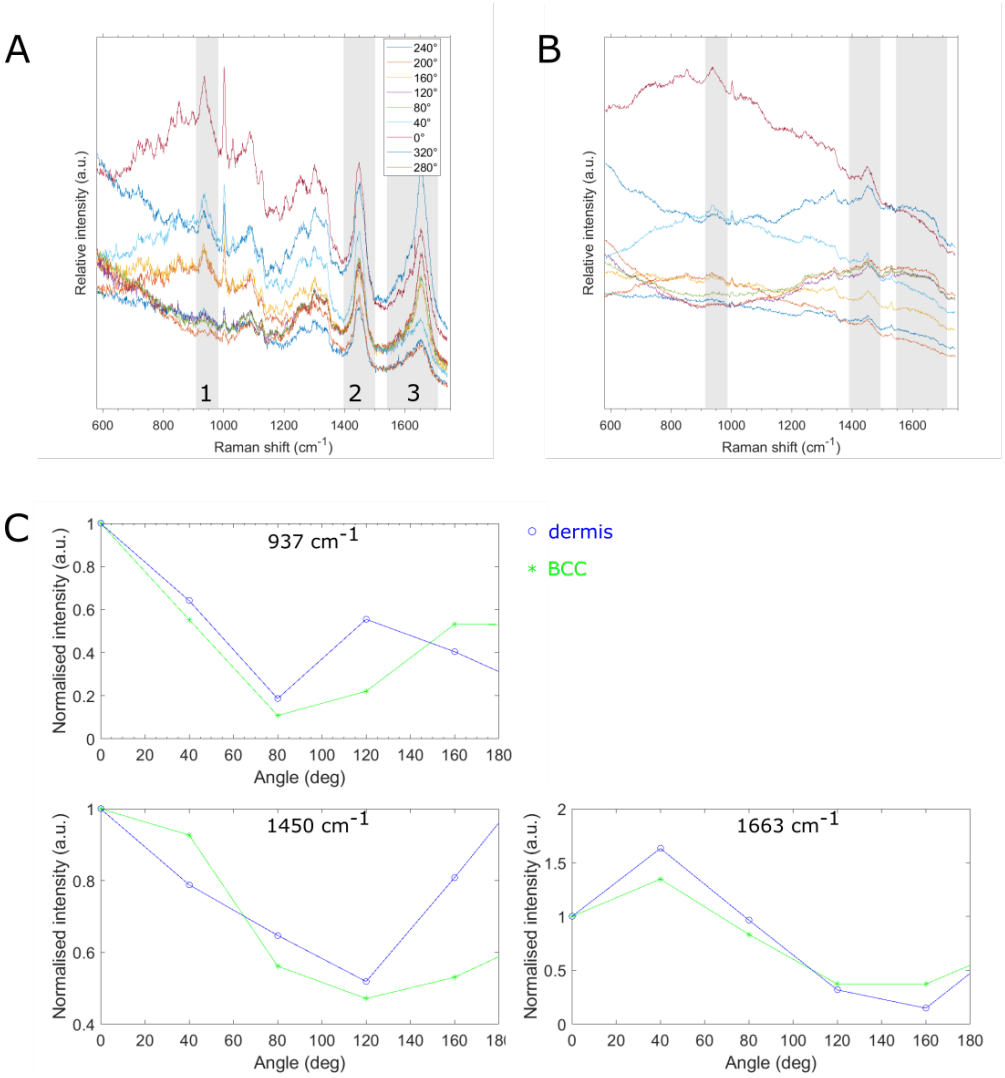


Figure 4.8: Polarised Raman spectroscopy at various angles of collection including parallel and perpendicular for a homogeneous tissue sample presenting BCC in **A** and a homogeneous sample of healthy dermis in **B**. In **C** are presented the depolarisation of the bands at 937 cm^{-1}, 1450 cm^{-1} and 1663 cm^{-1} for BCC (green) and dermis (blue) as a function of the angle θ with intensity normalised to the intensity in parallel configuration ($\theta = 0°$). They correspond to the bands shaded in grey and numbered as 1, 2 and 3 in **A**.

before. Thus, we decided to measure polarised Raman maps of skin tissues in parallel, at $\theta = 40°$ and in perpendicular.

Polarised Raman scanning at 0°, 40° and 90°

Figure 4.10 presents an example of the polarisation dependence study for a tissue with healthy stroma, sebaceous gland and inflamed regions.

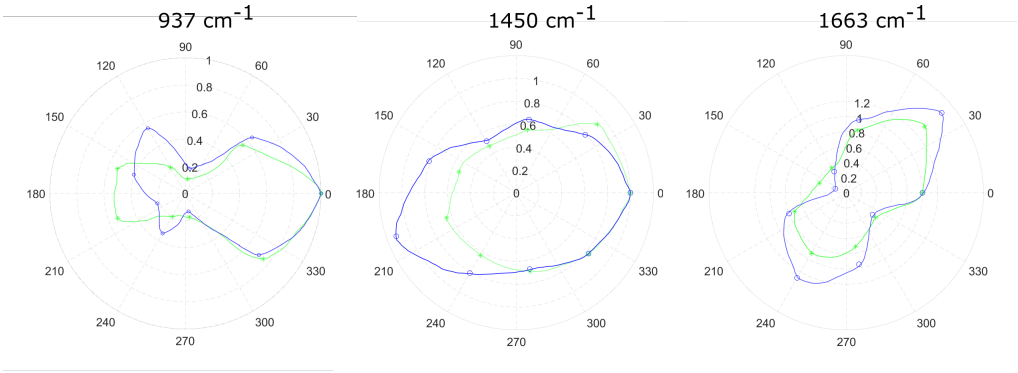


Figure 4.9: Polar diagrams of the peak intensity of the bands at 937 cm^{-1}, 1450 cm^{-1} and 1663 cm^{-1}. The data represented in blue correspond to the healthy dermis while the data points in green represent the basal cell carcinoma sample.

The excellent correlation between the PCA images and the H&E image, with recognizable features such as fat from the sebaceous gland, was the evidence of the reliability of the measurements. In the three angles of collection, there was exact overlapping of the PCA based images proving robust reproducibility of stage movement and that no degradation of skin tissue is happening with laser exposure (time and power). The raw spectra obtained present different level of fluorescence depending on the angle of collection generating variations of intensity of the Raman peaks with the angle. As mentioned in previous results, there is a loss of Raman information in the region below 1003 cm^{-1} for $\theta = 90°$, however the 40° configuration has rich information in this same region.

In the case of dermis, from region 1, the peak at 1663 cm^{-1} had a comparable intensity to the one at 1450 cm^{-1} for $\theta = 40°$ and 90° as expected from literature [53] however at $\theta = 0°$ there was a decrease and distortion of the broad peak 1550 - 1710 cm^{-1}. The spectra at $\theta = 40°$ differed from the one at $\theta = 0°$ with increased intensity for the RNA peak at 813 cm^{-1} and for the peak at 1065 cm^{-1}. Besides, the relative intensity of the band between 1020 – 1139 cm^{-1} (including assignment for phenylalanine of collagen and lipid) was higher for this configuration. On the other hand, the hydroxyproline band was stronger at 873 cm^{-1} and the DNA/RNA ring breathing modes at 1373 cm^{-1} were more pronounced in the parallel configuration. In the region 1, we noticed peak contribution from cotton for the peaks at 1337 cm^{-1} and 1528 cm^{-1}.

In the case of inflammation (region 2), peak at 851 cm^{-1} (vibration of amino acids, valine, tyrosine and polysaccharides) and at 1061 cm^{-1} (C-C skeletal stretching)

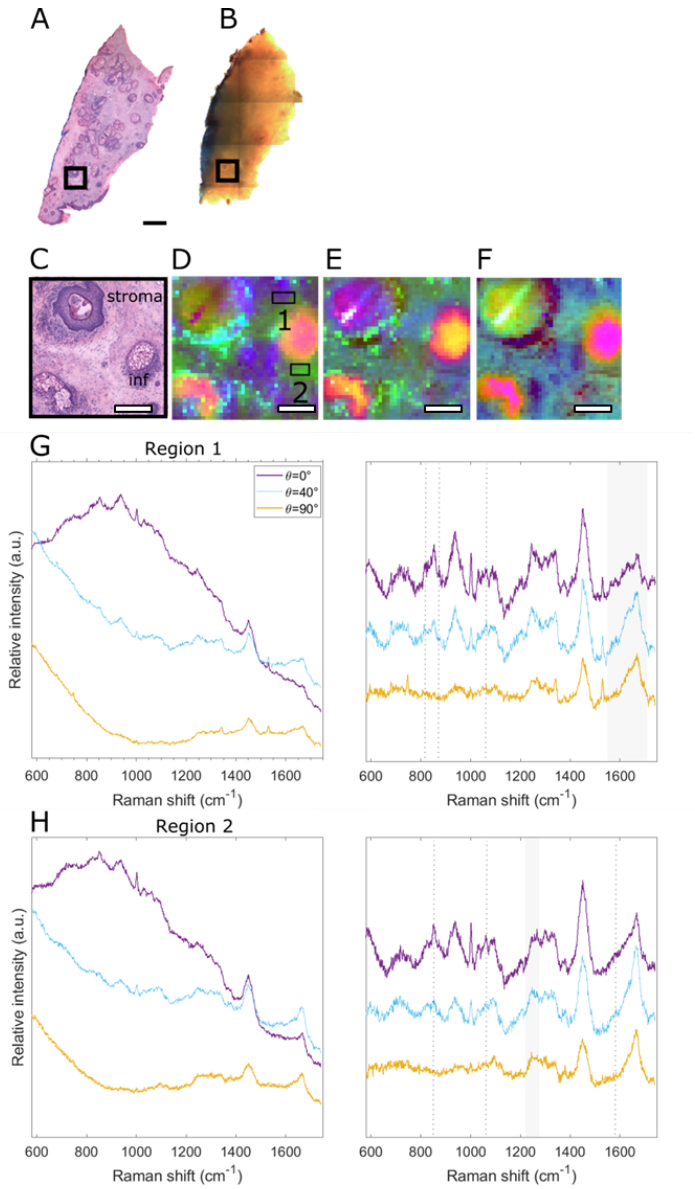


Figure 4.10: Polarised Raman spectroscopy of a skin tissue sample for three angles of collection: in parallel, at $\theta = 40°$ and in perpendicular.
We recorded microscopic bright-field images of the entire samples in **B**. The corresponding H&E images in **A** provided a detailed diagnosis of the tissue and allowed us to select areas of interest as represented by the black rectangle. The scale represents 1 mm. The PCA images obtained are displayed in **D** for parallel configuration, in **E** for the 40° angle collection and in **F** for perpendicular. They are on the same scale as the H&E region selected displayed in **C** (scale bar indicates 0.25 mm). We present in **G** the raw and baselined spectra from region 1 and in **H** the raw and baselined spectra from region 2. Inf stands for inflammation.

were stronger in the parallel-polarised spectra. Compared to normal dermis, the broad peak at 1550 cm^{-1} to 1710 cm^{-1} was as described in literature however it was only 84 % of the intensity of peak at 1448 cm^{-1}. The 40° collection angle

revealed the peak from α-helix and β-structure of the amide I band at 1637 cm^{-1} more than at $\theta = 0°$. In general, additional to these considerations, the slope was steeper in the 1200 cm^{-1} to 1250 cm^{-1} region (amide III) for dermis compared to inflammation which can be used for discrimination.

We repeated the experiment on a skin tissue with a different structure comprising basal cell carcinoma, stroma and epidermis, the results are presented in Figure 4.11.

We observed for spectra of normal dermis (region 1), the RNA peak at 813 cm^{-1} for $\theta = 0°$ and $\theta = 40°$. The hydroxyproline at 873 cm^{-1} was stronger for $\theta = 0°$ and the peak at 1086 cm^{-1} from ν(C-C) vibrations was more intense at 40° as described for the previous tissue. The broad peak at 1550 - 1710 cm^{-1} behaved in the same way. The region 2 representing the epidermis had specific peaks present for all configurations, such as 1301 cm^{-1} from CH_2 twisting, C-H vibration, $\tau(CH_2)$ of lipids. The vibration at 1057 cm^{-1} was more intense in parallel configuration while the peak from amide III band at 1245 cm^{-1} and at 1511 cm^{-1} from RNA were stronger for $\theta = 40°$. The region containing basal cell carcinoma (region 3) showed spectra representative of this tissue type with some distinct differences for the spectra at $\theta = 40°$ compared to $\theta = 0°$. The peak from C-C stretching of proline, valine and protein backbone at 935 cm^{-1} was more intense in parallel along with the peak at 1155 cm^{-1} from C-C(N) stretching of proteins.

In conclusion, to our knowledge, this is the first study reporting the use of polarised Raman mapping of skin tissue for a different configuration than the conventional parallel and perpendicular. The 40° angle configuration has the ability to reduce the collection of fluorescence while conserving spectral information relevant for diagnosis applications. These results will be further studied in section 4.4.

The following section describes the use of principal component analysis to visually differentiate the tissue types depending on their molecular differences.

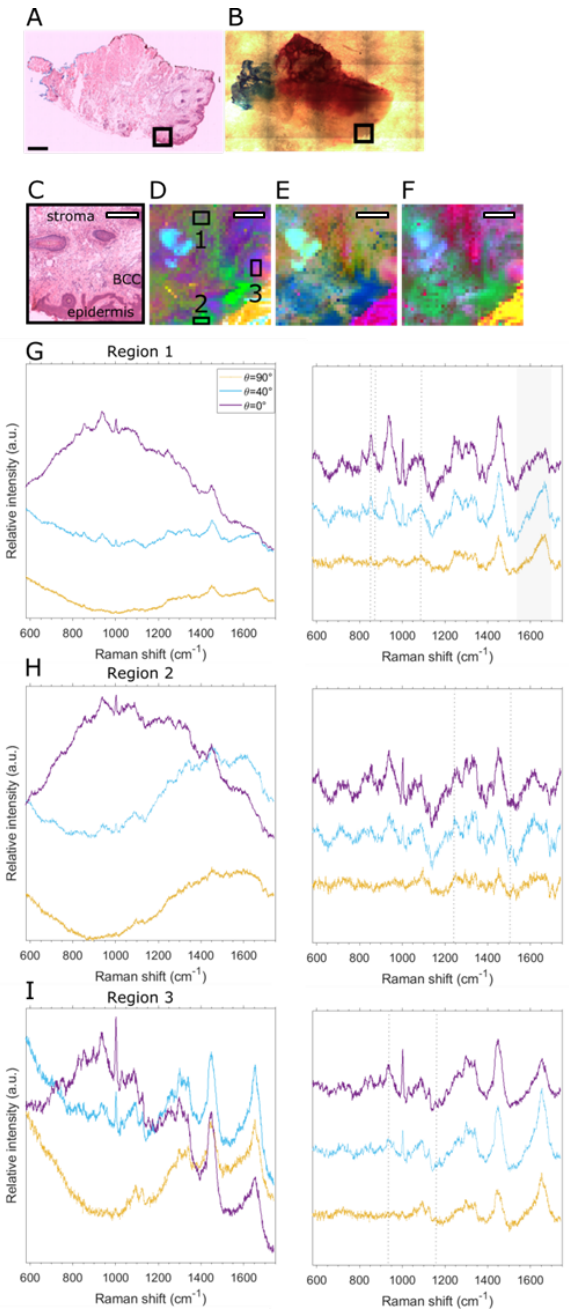


Figure 4.11: Polarised Raman spectroscopy of a skin tissue sample for three angles of collection: in parallel, at $\theta = 40°$ and in perpendicular.
We recorded microscopic bright-field images of the entire sample in **B** and of the corresponding H&E image in **A**. The scale represents 1 mm. The PCA images obtained are displayed in **D** for parallel configuration, in **E** for the 40° angle collection and in **F** for perpendicular. They are on the same scale as the H&E region selected displayed in **C** (scale bar indicates 0.25 mm). We present in **G** the raw and baselined spectra from region 1, in **H** raw and baselined spectra from region 2 and in **I** the raw and baselined spectra from region 3.

Principal component analysis (PCA)

In total, 13 samples have been measured in parallel and perpendicular polarised for the depolarisation study presented above. The average spectra for each type carried important Raman information that can be displayed with PCA scatter plots of the data [116]. Firstly, we performed PCA analysis on two separate set of data, on one hand, the spectra from the parallel angle collection and on the other hand the perpendicular angle of collection as shown in Figure 4.12. For both configurations, the analysis resulted in data points scattered over the entire 2D space with no cluster distinction. We observed that by no means differentiation was achieved using this method.

The approach proposed to solve the lack of discrimination between the tissue types was to combine the spectra in both configurations in a single vector. For each type, the co-polarised averaged spectra were baselined using a fourth order polynomial fit and the bisquare weight method over the full wavenumber range giving the vector $A = a_{parallel}$. In contrast, the cross-polarised spectra were reduced to the wavenumber range that contains the most Raman information between 900 cm^{-1} and 1750 cm^{-1} and baselined using a similar fifth order polynomial fit giving $B = a_{perpendicular}$. The final vector $C = [A \quad B]$ represented in Figure 4.13 was subjected to PCA analysis. The order of the polynomial fitting was determined considering the goodness-of-fit indicator: the coefficient of determination. This coefficient, R^2, tells how well the regression predictions approximate the real data points. We tested fourth to sixth-order polynomials fit, the best polynomial fit order was choosen to be optimised for each configuration maximising the R^2 coefficient.

Using this method of combination of spectra, we account for spectral variations of both angles of collection simultaneously. The principal component analysis can provide further information about the source of spectra variability and can be used as a differentiation tool of the various tissues' structures. The results of the analysis are presented in Figure 4.14.

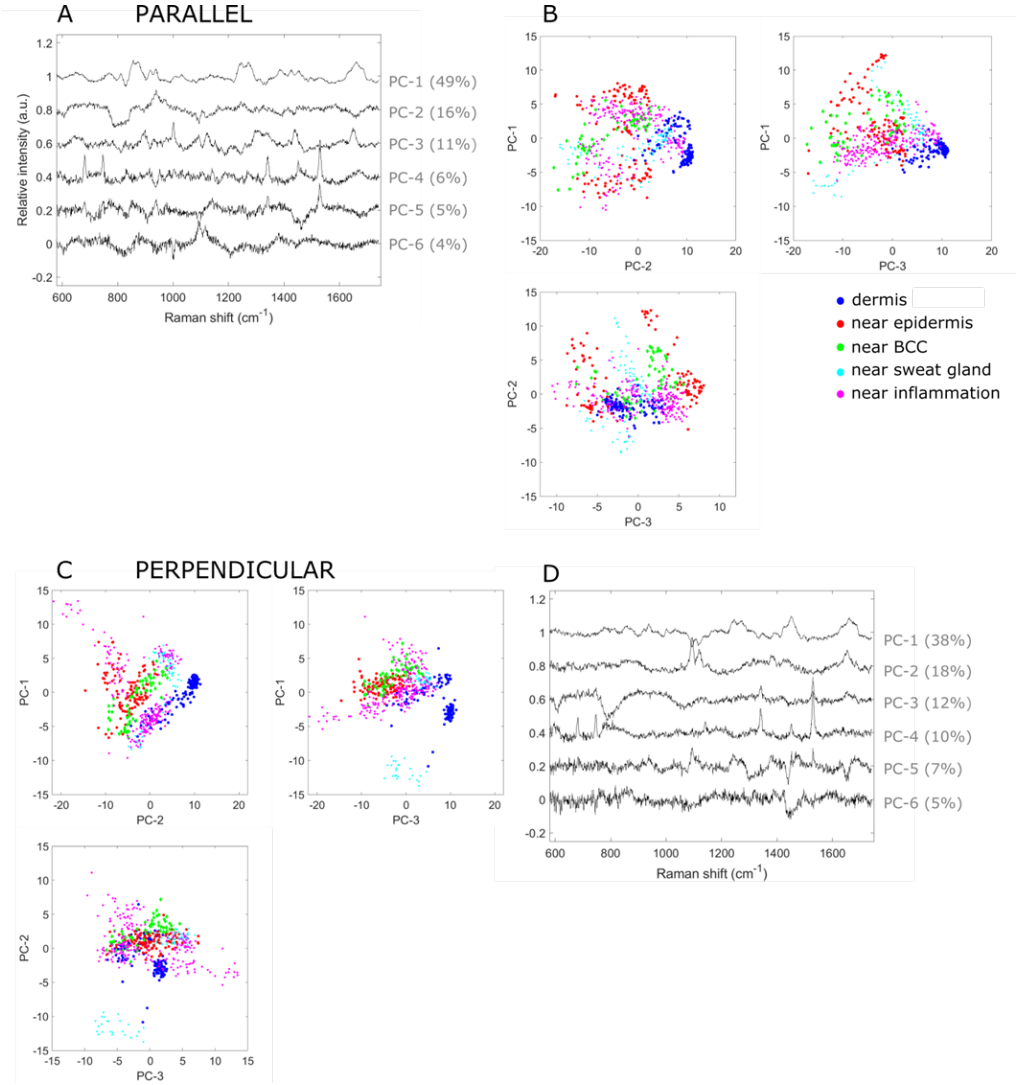


Figure 4.12: Plot of the loadings of the components PC-1 to PC-6 (offset for clarity) with corresponding percent variability for spectra in parallel configuration in **A**. Scores plots of different principal components combination after PCA was performed on the database in **B**. Similar plot of loadings for the perpendicular configuration is represented in **C** with scores plots in **D**. The colour code is as follow: dermis in blue, dermis next to epidermis in red, dermis next to BCC in green, dermis next to sebaceous gland in cyan and dermis next to inflamed tissue in magenta.

The spectral data points from dermis alone are grouped in a cloud of points separated from the other group. It means that the group from dermis alone is well discriminated from the other groups using the PCA method where PC-1 accounts for 95 % of the variance and PC-5 for 15 % of the variance. The loading of PC-1 contain Raman bands that can be assigned to molecular vibrations of amino acids such as proline, hydroxyproline and tyrosine (815 cm^{-1}, 857 cm^{-1}, 918 cm^{-1} and 950 cm^{-1}), tryptophan (872 cm^{-1}), to vibrations of protein such as

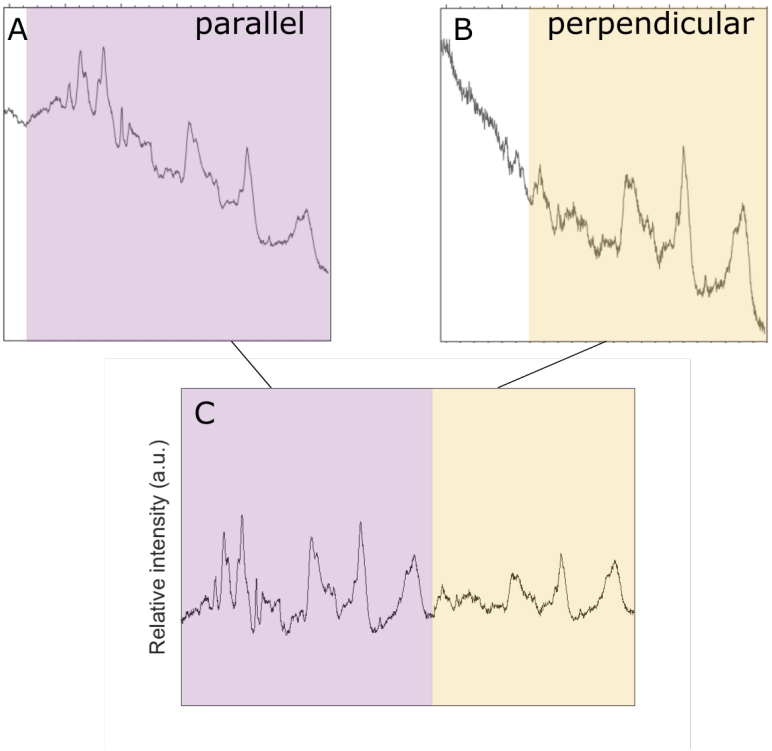


Figure 4.13: Illustration of the process of combining spectra in parallel configuration $a_{parallel}$ and in perpendicular $a_{perpendicular}$ configurations in one normalised unique vector C such as C = [$a_{parallel}$ $a_{perpendicular}$] for PCA analysis. Average spectra of dermis in co-polarised in **A** (area shaded in purple) is adjoined to the spectra in cross-polarised from the region with the most Raman information (area shaded in yellow) displayed in **B** to create the final vector C shown in **C**.

amide III of collagen (1249 cm^{-1}), amide III (1270 cm^{-1}), amide I (1665 cm^{-1}), phenylalanine (1003 cm^{-1} and 1033 cm^{-1}) and CH_2 bending (1450) [ref]. Even if PC-5 is a noisier spectrum, we can extract peaks from protein such as phenylalanine (1003 cm^{-1}), amide I (1665 cm^{-1} and 1650 cm^{-1}), tryptophan (1339 cm^{-1}) and collagen (1339 cm^{-1}), from lipids such as CH_2CH_3 deformation (1460 cm^{-1}), acyl backbone (1124 cm^{-1}), 1339 cm^{-1} and 1460 cm^{-1}. These results showed that the differentiation achieved reflect a combination of spectral signatures from many biochemical components ranging from protein, amino acids and lipids.

We observed that the data points were spread from the same tissue type groups in the scores plot from one sample to another illustrating inter-patient variability. However, each group contains between 70 and 220 spectra, so the groups were not totally equally represented which could affect the score plot examination. We also noted that the addition of the contribution from the perpendicular spectra

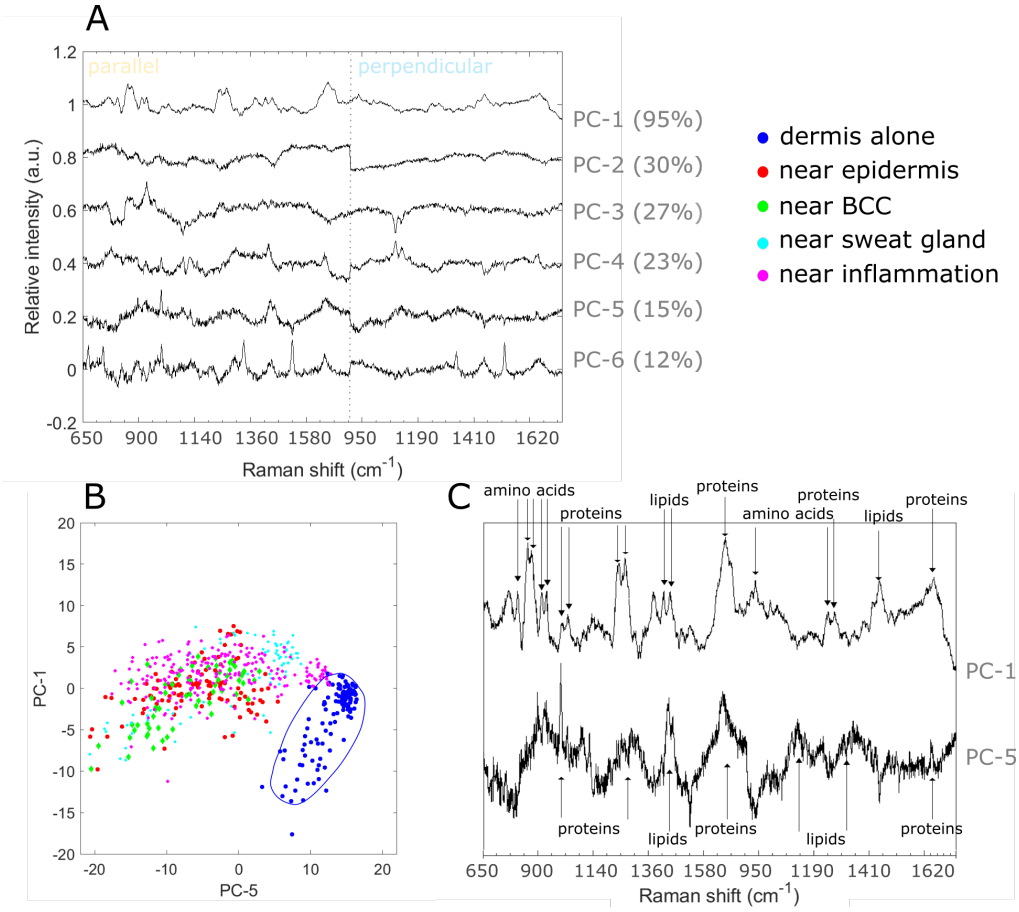


Figure 4.14: **A** Plot of the loadings of the components PC-1 to PC-6, the loadings have been offset for clarity. The dashed vertical line delineates the region where the spectra from parallel (left) and perpendicular (right) are combined. **B** Score plots of the two principal components PC-1 and PC-5 after PCA was performed on the thirteen samples represented by individual data points of different colours for healthy dermis asterisk (dark blue), and red, green, cyan and magenta respectively for dermis close to epidermis, BCC, sweat gland and inflamed region. **C** PC-1 and PC-5 loadings with peaks assignments.

was essential to differentiate the dermis tissue from the other tissues since many bands in the PC-5 loadings were identified from the part of the spectra originating from the perpendicular configuration.

Finally, the results indicate that skin is a complex tissue constituted of numerous biomolecular components and that healthy connective tissue has a specific composition which is affected by its vicinity with other molecular structures such as BCC or inflammation.

Discussion

Skin tissue is a turbid tissue, it scatters and diffuses light that travels through it. The scatter from the tissue comes from cells, vesicles, collagen fibrils, aggregates, membranes. Notably, skin dermis is a connective tissue composed of fibre structures. Even though in skin the collagen is not as orderly distributed as in tendon, it still exists in an organised structural arrangement [117]. However, the rapid depolarisation of linear polarised incident light was measured in skin tissue due to collagen fibres not perfectly aligned. On the other side, epidermis was showing little depolarisation [118]. This result highlights the effect of the depth of laser light penetration on the depolarisation of skin layers. We maintained during the entire study a similar laser depth penetration into the skin tissues to avoid this dependence.

A study on collagen-rich region of human skin tissue discussed how the band at 939 cm^{-1}, from C-C stretch, nearly vanishes when the sample is oriented perpendicular to the polarisation of the incident laser [118]. It could explain our observations in the lower wavenumber region (<1004 cm^{-1}) where we generally observed for almost all tissue types the disappearance of the Raman bands in this region.

The bands at 856 cm^{-1} and at 936 cm^{-1} are highly polarised ($\rho_{856,936} < 0.75$). This strong polarization could be coming from groups of molecules for which the change in polarizability has high spherical symmetry [119]. Also, a study on rat tendon tail showed the anisotropy of the Raman band at 1665 cm^{-1} with a higher intensity of this amide I band in the perpendicular direction to the incident laser beam when the collagen fibrils were oriented in the parallel direction [52]. The conclusion is confirmed by the human skin study mentioned earlier [118]. It emphasizes the fact that skin tissue is composed of molecules with an organisational structure and that amide I band difference in the two different polarisation configurations is due to the preferential orientation of collagen fibre in the skin tissue along one direction.

The depolarisation ratio analysis for skin tissues reported in literature include depolarisation ratio of bands at 623 cm^{-1}, 920 cm^{-1}, 940 cm^{-1}, 1065 cm^{-1}, 1128 cm^{-1}, and 1339 cm^{-1} [120]. The values for bands at 920 cm^{-1} and 940 cm^{-1} were

measured to be different for peritumoral stroma and healthy dermis, in our study we observed that the higher value for the band belong to healthy dermis. Our results are in accordance with the literature. Another study was interested in the depolarisation ratio of the bands at 920 cm^{-1}, 940 cm^{-1}, 1065 cm^{-1}, 1128 cm^{-1}, 1210 cm^{-1}, and 1247 cm^{-1} showing depolarisation for the bands at 920 cm^{-1} and 940 cm^{-1} to be lower than 0.75 matching our observation for the band at 937 cm^{-1} however it is unclear how they separate the bands at 920 cm^{-1} from the band at 940 cm^{-1} since they are overlapping in the range between 900 cm^{-1} - 990 cm^{-1} [60].

Finally, we conducted a preliminary investigation to reconfirm the potential of the 40° angle orientation to attenuate collection of the fluorescence background from highly fluorescent tissues.

4.4 Application on fluorescent tissues

Introduction

Fluorescence is a competing process likely to interfere with the Raman signal and in certain cases, fluorescence makes the interpretation of the scattered signal difficult [121]. The incident photons absorbed experience internal conversion and are re-emitted at a longer wavelength with a higher probability compared to Raman scattering. It is due to a quantum yield 6 to 10 orders of magnitude higher than the one from Raman signal. In biological tissues, fluorescence is induced by intrinsic fluorophores such as collagen, elastin, melanin for skin [122], tryptophan, melanin, kynurenine for hair [123], elastin, collagen, NADH, haemoglobin for lung [36, 73, 124]. The fluorescence signal induced by the fluorophores is a drawback in Raman measurements, its contribution creates an intense fluorescence background which is most of the time hard to separate from the Raman signal. In some cases, the fluorescence is saturating the detector and even a longer integration time would have no effect. Solutions to help

to remove fluorescent background were mentioned, they can be experimental techniques such as a shift of the incident excitation wavelength to 1064 nm, a higher wavelength [77, 125], temporal gating [77], photobleaching, or mathematical techniques like subtraction algorithm [126], deconvolution techniques and polynomial curve-fitting. Although some of the methods mentioned to suppress fluorescence background have been proposed, some imply the purchase of expensive additional components or the development of a totally new setup which could be avoided. Also, to our knowledge, only a few methods have shown interest in utilizing the propriety of polarization for reducing fluorescence [23]. In this part of the study, we tested our hypothesis that the fluorescence emission from a biological sample might have different polarisation characteristics to the Raman scattered light depending on the orientation of its molecular components. Therefore, the choice of the correct direction of collection could lead to the collection of spectra with reduced fluorescent background while conserving the Raman signal. We were interested in tissues with low Raman signal and high fluorescence such as liver [127], kidney [128] and lung.

Material and methods

Lung tissue samples were obtained from surgery at the Nottingham University Hospital NHS Trust, City Campus with ethical approval from the Nottingham Health Science Biobank. Lamb kidney and pork liver were obtained as fresh samples from a local supermarket. Tissues were defrosted at ambient temperature, washed in PBS and placed on a quartz slide (100 μm thick) for measurement. Spectra, in all configurations mentioned, were recorded at each position in the fingerprint region (700 - 1750 cm^{-1}) with an acquisition time of 2 s per spectrum. To minimize any possible damage to the tissues with time, data presented here were acquired within a 6-hour period.

Lung tissues

We present in Figure 4.15, the Raman spectra collected from lung tissues along four directions compared to the laser incident polarisation including parallel and

perpendicular directions. It must be emphasized that two homogenous samples were chosen for this experiment. The four directions presented were selected for their interesting features following the study of multiple angles presented in 4.3. The degree indicates the analyser direction compared to the laser polarisation which means that parallel configuration is represented by $\theta = 0^\circ$.
Parallel-polarised spectra of lung tissues showed that healthy lung tissue has a lower level of fluorescence background compared to lung tumour. We observed that the direction of polarisation collection influences the level of fluorescence background in both cases. Indeed, for 40° and 200° the fluorescence drastically dropped and seemed to reveal some Raman peaks with a higher intensity over the intense fluorescence background. This effect was visible for healthy lung tissue and particularly for lung tumour tissue for which the detection of Raman signal at 40° particularly enhances the protein Raman peak at 1655 cm^{-1}. This peak, corresponding to the C=O stretching of elastin and collagen and the amide I band of protein in the α-helix conformation [36] has non-negligible importance for diagnosis purposes. Our observations of the perpendicular-polarised spectra for lung tissues for normal and tumour samples had similar features to skin tissues described in 4.3.

Indeed, we noticed a decrease of the fluorescence background but mainly the disappearance of Raman bands in the 600 cm^{-1} to 1000 cm^{-1} spectral range. The bands in this region are a rich source of information for discriminating tissue types since they are the results of more than 150 vibrations of molecules [92] ranging from protein vibrations C-C stretching mode of proline, valine, α-helix of protein backbone, α-helix of keratin, ν(C-C), tyrosine to polysaccharide vibrations and collagen.

Other tissues

We also conducted measurements on two known highly fluorescent tissues: liver and kidney which we present in Figure 4.16.

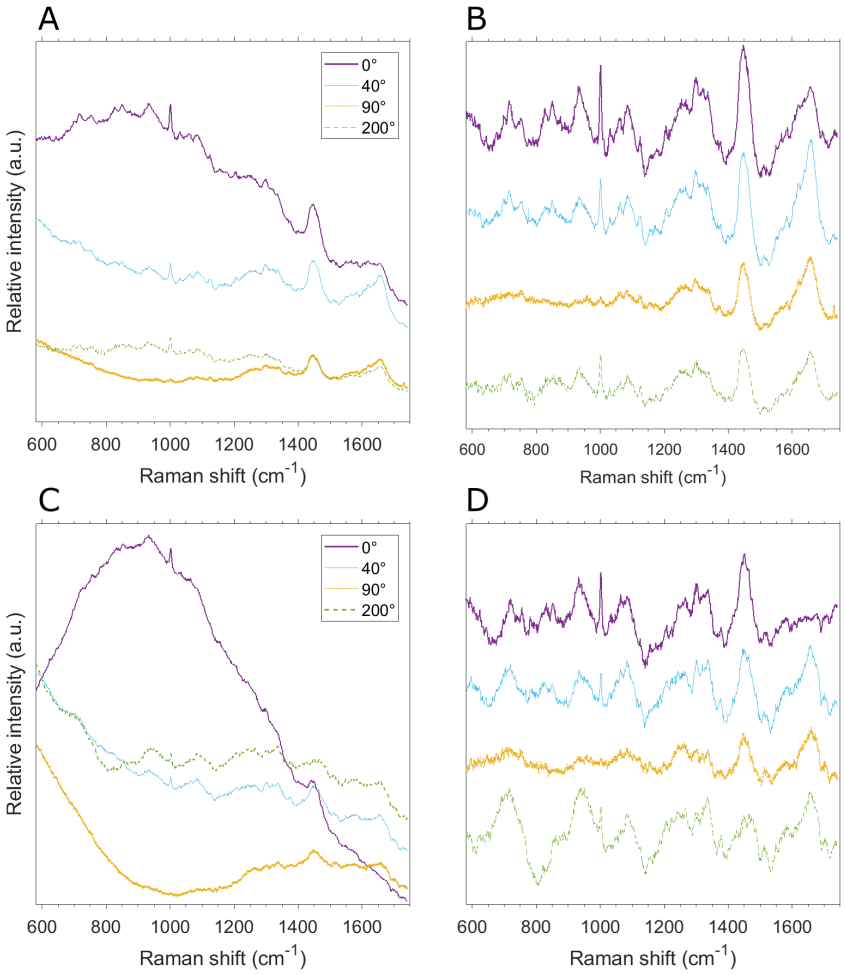


Figure 4.15: Dependence of Raman spectra with the polarisation angle collection (θ). Polarised Raman spectra of tissues at θ = 0°, 40°, 90° and 200° recorded for healthy lung tissues in **A** and lung malignant tissues in **C**. Spectra in **A** and **C** are background corrected. They are presented with corresponding fifth-order polynomial subtracted spectra in **B** and **D**.

In the fingerprint region, the non-processed spectra were dominated by a main fluorescence peak mostly observed in parallel configuration. Once again, the fluorescence peak at 900 cm^{-1} was largely dropping with the 40° collection. For lamb tissue, bands were revealed at 1104 cm^{-1} from phenylalanine and between 1618 cm^{-1} and 1668 cm^{-1} with the appearance of a band at 1649 cm^{-1} from the C=C vibrations of amide I. For kidney tissue, the band between 1620 cm^{-1} and 1672 cm^{-1} was enhanced for the 40° configuration and overall for both tissues, the parallel configuration revealed peaks in the region 1202 – 1280 cm^{-1} not observed for θ = 40°.

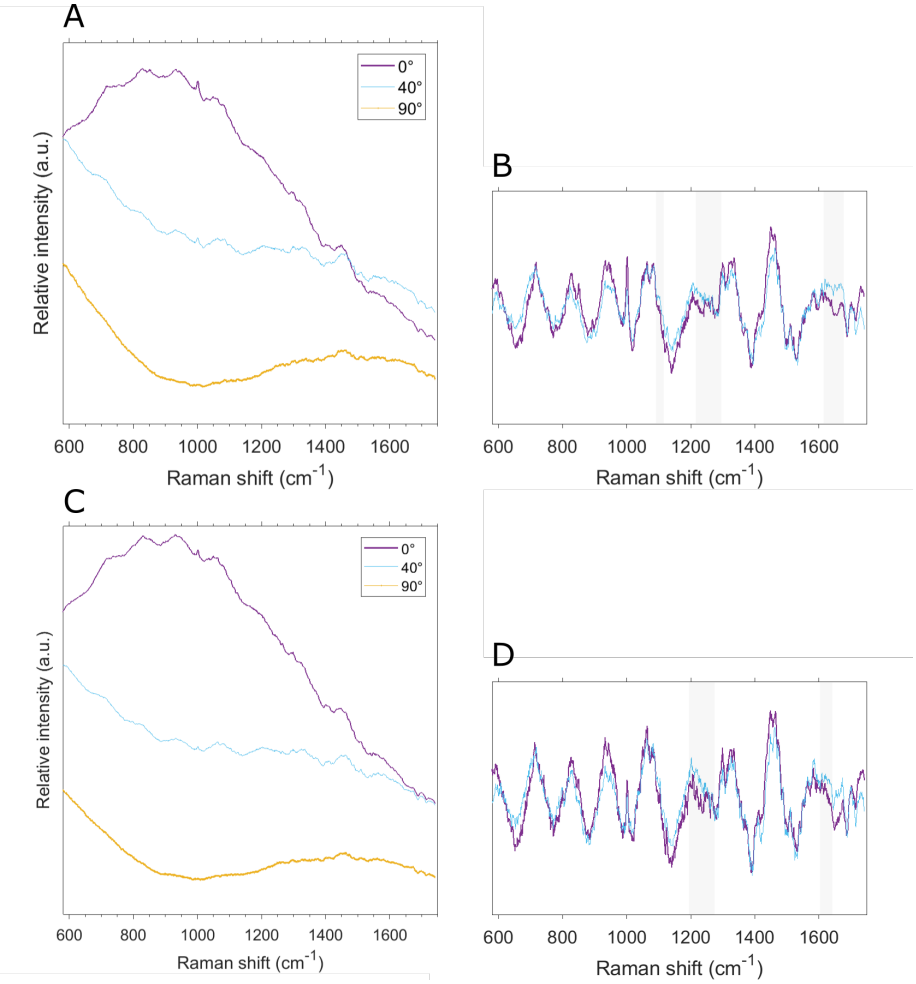


Figure 4.16: Raw data averaged from 100 spectra from polarised Raman spectroscopy measurement of lamb kidney in **A** and pork liver in **C** with their corresponding baseline-corrected spectra in **B** and **D** at $\theta = 0°$ and $\theta = 40°$.

Discussion

Biological tissues contain densely packed interconnected scatterers in a fixed matrix. We can expect from the regular arrangement of the scatterers in a tissue, a high level of birefringence which would lead to a loss in the polarisation state of the incident excitation. Studies on transverse cross section of porcine adipose and myocardial tissue of pigs show that tissues with no birefringence lead to the conservation of the polarization state of light incident on the tissue and the polarization state of light back-scattered from it [129]. Comparable results have been observed in chicken tissue [118] where muscle display birefringence from actin-myosin filaments on the contrary to liver which has negligible birefringent

fibres. The work mentioned suggests a rapid randomization (i.e. depolarisation) of linear polarisation in birefringent tissues. The oriented collagen in fibres rapidly depolarised the linear polarised incident light hitting the skin whereas for example chicken liver, will exhibit a slow depolarisation. In addition to these considerations, the study from Studinski et al. [130] reveals an interesting aspect: polarisation measurement in backscattered configuration has unique preservation properties in tissue-like media and it might also be true in biological tissue. Besides, biological tissues are said to not only be circularly birefringent but also multi-scattering tissues and likely to present direction-dependent birefringence. Smith et al. [119] explained that there is retention of polarisation state for single Raman scattering events for randomly oriented molecules in a tissue, with no dependence on the Raman mode and when the polarizability of a molecule has high spherical symmetry. It means that multiple scattering events will randomize both the direction and state of propagation of light [131]. Furthermore, white-light study on skin tissues from Mohs surgery showed that only three per cent of the incident light is retaining linear polarisation whereas around 93 % of light becomes randomly depolarised [118]. It is due to light penetrating deep in the tissue which is responsible for the occurrence of multiple scattering events. In terms of polarisation proprieties of fluorescence, the time delay between the absorption of a photon and its re-emission leaves space for the excited molecule to disorient leading in a reduction of the polarisation [119]. As we have concluded that biological tissue investigated in our system configuration contain randomly oriented molecules, it is expected that fluorescence and Raman scattering will have different behaviour in terms of polarisation and depolarisation signals. This concept was observed in the case of chemicals on quartz [23], where the author, using the depolarised nature of fluorescence and Raman signal, obtained a signal with reduced fluorescence in separating the depolarised fluorescence background from the Raman peaks of chemical of interest.

In our case study, we are presenting a preliminary investigation showing that for human lung, lamb kidney and pork liver tissue samples the selection of a specific angle reduces the collection of fluorescence and that the remaining spectra have excellent spectral information for diagnostics applications. This observation reveals that fluorescence emission is likely to have preferential directions of

emission and that these directions differ to the ones from the Raman signal. These experiments also confirm our previous observations on skin tissues revealing that the perpendicular-polarised spectra present poor molecular information.

In summary, Raman mapping polarised measurements were obtained for 15 skin tissue samples from 7 patients. The maps showed a strong polarisation dependence of many Raman bands. The calculation of depolarisation ratios showed, among others, that there was a high level of alignment in nucleic and amino acids for healthy dermis not observed for peritumoral dermis, or dermis close to sweat gland or epidermis. Then, different angles for the collection of signal were investigated, including parallel, perpendicular, and 40°. The 90° was combined with the 0° collection. The combination of both contributions allowed successfully to discriminate the dermis from healthy tissue and the dermis at the vicinity of other molecular structures, with the PCA method. The results emphasize the conclusion that Raman polarised spectroscopy is suitable for the characterisation of inherent morphological structure in biological tissues. Finally, this polarised technique was applied to highly known fluorescent tissues. The collection of fluorescence background was drastically reduced for example for lung tissues at specific angles. We concluded that the fluorescence emission has a direction of emission different from the one from the Raman signal making the polarisation properties relevant for this application.

5 Conclusion

5.1 General conclusions

In **chapter 3**, we described the *ex vivo* measurement of lung tissues obtained with the confocal Raman microspectrometer. The Raman mapping measurements of the tissues presented allowed, in combination with histopathology staining, to provide a detailed molecular composition of lung tissues. By analysing the Raman spectra obtained from malignant and non-malignant fresh lung tissues, we have demonstrated the possibility of discrimination between malignant and healthy tissues. We showed that specific ratio of intensity and spectral differences are potential markers for discriminating adenocarcinoma, squamous cell cancer and normal tissues. The results also showed that fluorophore environment variations such as high collagen content can explain the patient intra and inter variability in terms of fluorescence levels. Besides, we associated the variation in concentration of porphyrin with the stage of the necrotic tissue.

Our results are based on mapping measurements with high spatial resolution and accurate correlation with histopathology diagnosis. Consequently, the k-means images revealed the tissue heterogeneity highlighting Raman spectral characteristics at the micron scale. In various studies, the fluorescence background from the lung tissue made the Raman experiments challenging. In this study, we gave a clear illustration of the feasibility to discriminate lung tissue types towards a diagnosis application. We highlighted the common challenges encountered by working with fresh human lung tissues, and we are the first to report their occurrence. The Raman mapping measurements provide precise, label-free and rapid chemical information. It suggests that it could be applied to reduce the percentage of unclassified or poorly differentiated tumours. In fact, the Not Otherwise Specified (NOS) type reported in approximately 10 per cent of the non-small

lung cancer cases requires an improvement in the tumour subtyping. It is relevant in the context of emerging treatment developed with specific therapeutic agents. Therefore, Raman measurements have high potential in speeding the diagnosis and gathering spectral information regarding cancer types and subtypes aiding the histopathologist in identifying biomolecular changes with objectivity. This study supports the use of Raman spectroscopy as a future method supplementing histopathology in clinical practice. Finally, the study conclusions emphasise the possible future application of Raman spectroscopy to complement the current diagnostic methods available *in vivo* and contribute to the understanding of the chemical mechanisms involved in cancer development, tissue necrosis, and cell local environment.

The results presented in **chapter 4** described polarised Raman mapping spectroscopic investigation of skin structures guided by H&E stained images. High-quality maps were obtained of skin regions which provided information about their molecular composition and alignment.

Skin connective tissue was characterised by polarised Raman measurements revealing that Raman bands have a strong polarization dependence. Indeed, the parallel angle collection showed variations in band intensity compared to the perpendicular collection. Overall, the cross-polarised collected spectra were not as rich in spectral information as the co-polarised spectra however the combination of both brought insightful molecular details. The depolarisation ratio calculations allowed the discrimination of the dermis as a function of its local environment, suggesting collagen modification (band at 937 cm^{-1}), nucleic and amino acid conformation variation (band at 856 cm^{-1}) and different molecular alignment of lipids (band at 1450 cm^{-1}) for dermis alone compare to dermis in the proximity of BCC, epidermis and inflammation. The angle dependency measurements revealed preferential orientations for molecules within tumorous and non-tumorous tissues not reported in the literature before such as a preferential polarisability of groups of molecules responsible for the band at 937 cm^{-1} tilted compared to the group of molecules at 1663 cm^{-1}. This part of the study brought our attention to the potential of the 40° angle configuration in reducing the fluorescence background collection while maintaining a level of Raman information that can be used for

tissue diagnosis such as ratio of band at 1312 cm^{-1} and 1268 cm^{-1} [109]. The PCA analysis of the spectra in both configurations combined confirmed the previous results indicating biomolecular variations in connective tissue within skin tissues, making the contribution of lipids and protein a potential discriminative marker.

5.2 Future directions

Our work on human lung tissue samples is part of the existing Raman spectroscopic studies. With the use of a microscope-based laboratory instrument and not a commercial spectrometer, we achieved major improvements in terms of data collection. First of all, the acquisition time was drastically reduced compared to all other studies. The measurement time of 1 s achieved here makes the application to this technique in clinical routine more realistic. Additionally, no sample preparation was needed to keep the tissues in their original physiological state, even though tissue metabolism should still be investigated to correct the translation of the results from *ex vivo* to *in vivo*. Second of all, we performed Raman mapping measurements directly correlated to the tissue H&E adjacent sections. This unique way of creating a spectral database is the most complete study, to our knowledge, reported in the literature. Finally, the discrimination of tissue cancer types by multivariate statistical analysis was applied highlighting the extensive list of molecular differences between them. On top of that, we are the first to report statistics on the occurrence of laser-induced damage and tissue shrinkage. The current work provides additional information to the study of lung cancer tissues by Raman spectroscopy and shows the viability of future application in cancer diagnosis. For the clinical translation of this study, there is clear scope for further refinement of the methodology such as increasing the size of the region measured in the Raman map. We also think that sample preparation can be improved for example by developing a consistent method for controlling the thickness of tissue samples and for the blood removal procedure. Besides, more research on the understanding of the source of fluorescence need to be pursued. Additionally, it would imply to work on a larger database and developing an algorithm of classification that will be trained on the database. Then, from

the microscope-based instrument, we should develop a prototype instrument based on the requirements and needs of the hospital staff and histopathologist. Considerations such as the space available for the prototype (design, dimension, packaging), its ease of use (automated or manual operating procedure), the expected performances (sensibility, specificity, SNR, measurement time) will be largely discussed. For example, the miniaturization system introduced in previous studies [132, 133] could be considered. More importantly, joint discussions should define the aim of the prototype. Would the instrument help highlight regions of different tissue morphology in the tissue? Would it be more relevant to focus on the detection of a certain cancer type on a surgical margin? Should we add autofluorescence confocal imaging to the prototype to guide the Raman signal collection to specific points of interest?

Regarding the study based on polarised Raman spectroscopy, it is the first work to our knowledge to use polarised Raman mapping measurement as opposed to single point measurements of human skin tissues. This method means that the tissue sample can be investigated at both high axial and lateral resolution, two important parameters for the case of skin tissue since it has a highly varying molecular composition at a microscopic level. The observations reported were based on the average spectra from regions in the sample however we can envision using the information contained in the entire map to study the more subtle influences of the structures on the healthy dermis. For instance, by calculating the band area of a specific group of molecules and by representing a depolarisation ratio map, it would create a visual representation of the intensity of a band depending on its localisation in the tissue. Could we discriminate lung cancer types from band intensity maps? Or from the ratio of intensity of bands at every point in the map? Although, for now, the observations are not all well understood mainly because of unassigned bands; the study of depolarisation ratio map might be crucial for understanding the biochemical changes gradually happening within the skin tissue because of a structural particularity. Could polarised Raman mapping unravel biomolecular structural changes in skin cancer never observed by conventional Raman mapping? This could lead to the discovery of lipid or protein-specific organisation pattern. Also, as exposed in another polarised Raman

study [55], assessing the vibrations both parallel and perpendicular combined might reduce the false diagnosis results obtained from one configuration only. For example, the study of connective tissue could be envisioned. Indeed, the connective tissue is mainly composed of collagen which maintains structural integrity. Its degradation by matrix metalloproteinases (MMPs) is activated by metabolic stress along with tumorigenesis and inflammation [134]. The bands induced by collagen vibrations present in the Raman spectra of co-polarised and cross-polarised spectra could reflect such degradation which might serve as an important marker of tumour invasiveness. Such studies may benefit fundamental skin research in fields such as dermatology and metabolomics.

References

[1] Sheikh Mohammed Shariful Islam, Tina Dannemann Purnat, Nguyen Thi Anh Phuong, et al. "Non-Communicable Diseases (NCDs) in developing countries: a symposium report". *Globalization and health* 10.1 (2014), p. 81 (cit. on p. 1).

[2] John M Balbus, Robert Barouki, Linda S Birnbaum, et al. "Early-life prevention of non-communicable diseases". *The Lancet* 381.9860 (2013), pp. 3–4 (cit. on p. 1).

[3] RW Field, BJ Smith, CE Platz, et al. "Lung cancer histologic type in the surveillance, epidemiology, and end results registry versus independent review". *Journal of the National Cancer Institute* 96.14 (2004), pp. 1105–1107 (cit. on pp. 1, 54).

[4] John E Cotes, David J Chinn, and Martin R Miller. *Lung function: physiology, measurement and application in medicine*. John Wiley & Sons, 2009 (cit. on p. 4).

[5] Cancer research uk. *lung cancer*. 2016. URL: `https://www.cancerresearchuk.org/about-cancer/lung-cancer` (cit. on p. 5).

[6] Sarah Walters, Sara Benitez-Majano, Patrick Muller, et al. "Is England closing the international gap in cancer survival?" *British journal of cancer* 113.5 (2015), p. 848 (cit. on pp. 5, 54).

[7] Keith M Kerr. "Personalized medicine for lung cancer: new challenges for pathology". *Histopathology* 60.4 (2012), pp. 531–546 (cit. on p. 5).

[8] Lauren G Collins, Christopher Haines, Robert Perkel, and Robert E Enck. "Lung cancer: diagnosis and management". *Am Fam Physician* 75.1 (2007), pp. 56–63 (cit. on p. 6).

[9] James G Marks and Jeffrey J Miller. *Lookingbill and Marks' Principles of Dermatology E-Book*. Elsevier Health Sciences, 2017 (cit. on p. 7).

[10] Jane Sandby-Moller, Thomas Poulsen, and Hans Christian Wulf. "Epidermal thickness at different body sites: relationship to age, gender, pigmentation, blood content, skin type and smoking habits". *Acta Dermato Venereologica* 83.6 (2003), pp. 410–413 (cit. on p. 7).

[11] Karan Chopra, Daniel Calva, Michael Sosin, et al. "A comprehensive examination of topographic thickness of skin in the human face". *Aesthetic surgery journal* 35.8 (2015), pp. 1007–1013 (cit. on p. 7).

[12] M Ulrich, E Stockfleth, J Roewert-Huber, and S Astner. "Noninvasive diagnostic tools for nonmelanoma skin cancer". *British Journal of Dermatology* 157 (2007), pp. 56–58 (cit. on p. 9).

[13] Cristiano Rampinelli, Daniela Origgi, and Massimo Bellomi. "Low-dose CT: technique, reading methods and image interpretation". *Cancer Imaging* 12.3 (2012), p. 548 (cit. on p. 10).

[14] Phillip M Boiselle. "Computed tomography screening for lung cancer". *Jama* 309.11 (2013), pp. 1163–1170 (cit. on p. 10).

[15] Piergiacomo Calzavara-Pinton, Caterina Longo, Marina Venturini, Raffaella Sala, and Giovanni Pellacani. "Reflectance confocal microscopy for in vivo skin imaging". *Photochemistry and photobiology* 84.6 (2008), pp. 1421–1430 (cit. on pp. 11, 12).

[16] Andrew John Coleman, Thomas James Richardson, Guy Orchard, et al. "Histological correlates of optical coherence tomography in nonmelanoma skin cancer". *Skin Research and Technology* 19.1 (2013), e10–e19 (cit. on p. 13).

[17] Syril Keena T Que. "Research Techniques Made Simple: Noninvasive Imaging Technologies for the Delineation of Basal Cell Carcinomas". *The Journal of investigative dermatology* 136 (2016), e33–e38 (cit. on p. 13).

[18] G-Andre Banat, Aleksandra Tretyn, Soni Savai Pullamsetti, et al. “Immune and inflammatory cell composition of human lung cancer stroma”. *PLOS one* 10.9 (2015) (cit. on p. 14).

[19] Geoffrey D Rubin. “Lung nodule and cancer detection in CT screening”. *Journal of thoracic imaging* 30.2 (2015), p. 130 (cit. on p. 14).

[20] GO Rolls. “101 Steps to Better Histology—a Practical Guide to Good Histology Practice”. *Leica Microsystems* (2008), pp. 1–137 (cit. on p. 15).

[21] Martial Ducloy. “The laser, from yesterday (1960) to tomorrow”. *Europhysics News* 42.1 (2011), pp. 17–19 (cit. on p. 16).

[22] James R Janesick et al. *Scientific charge-coupled devices*. Vol. 117. SPIE press Bellingham, 2001 (cit. on p. 16).

[23] Steven Christesen. “Improved Raman sensitivity using polarization analysis”. In: *Chemical, Biological, Radiological, Nuclear, and Explosives (CBRNE) Sensing XI*. Vol. 7665. International Society for Optics and Photonics. 2010, 76651B (cit. on pp. 20, 111, 115).

[24] F Siebert and P Hildebrandt. “Theory of infrared absorption and Raman spectroscopy”. *Vibrational Spectroscopy in Life Science* (2008), pp. 11–61 (cit. on p. 21).

[25] Derek A Long. *The Raman Effect*. Wiley, 2002 (cit. on p. 22).

[26] Richard L McCreery. *Raman spectroscopy for chemical analysis*. Vol. 225. John Wiley & Sons, 2005 (cit. on pp. 22, 61).

[27] Rebecca Richards-Kortum and Eva Sevick-Muraca. “Quantitative optical spectroscopy for tissue diagnosis”. *Annual review of physical chemistry* 47.1 (1996), pp. 555–606 (cit. on p. 22).

[28] Georges A Wagnieres, Willem M Star, and Brian C Wilson. “In vivo fluorescence spectroscopy and imaging for oncological applications”. *Photochemistry and photobiology* 68.5 (1998), pp. 603–632 (cit. on p. 23).

[29] GC Tang, Asima Pradhan, Wenling Sha, et al. “Pulsed and cw laser fluorescence spectra from cancerous, normal, and chemically treated normal human breast and lung tissues”. *Applied optics* 28.12 (1989), pp. 2337–2342 (cit. on p. 23).

[30] RRTG Alfano, GC Tang, Asima Pradhan, et al. “Fluorescence spectra from cancerous and normal human breast and lung tissues”. *IEEE Journal of Quantum Electronics* 23.10 (1987), pp. 1806–1811 (cit. on p. 23).

[31] Christoph Krafft, Daniela Codrich, Gloria Pelizzo, and Valter Sergo. “Raman and FTIR imaging of lung tissue: Methodology for control samples”. *Vibrational Spectroscopy* 46.2 (2008), pp. 141–149 (cit. on pp. 23, 55, 63, 70, 71).

[32] Frederik Großerueschkamp, Angela Kallenbach-Thieltges, Thomas Behrens, et al. “Marker-free automated histopathological annotation of lung tumour subtypes by FTIR imaging”. *Analyst* 140.7 (2015), pp. 2114–2120 (cit. on p. 23).

[33] Xiaoliang Sun, Yizhuang Xu, Jinguang Wu, Yuanfu Zhang, and Kelin Sun. “Detection of lung cancer tissue by attenuated total reflection–Fourier transform infrared spectroscopy—a pilot study of 60 samples”. *Journal of Surgical Research* 179.1 (2013), pp. 33–38 (cit. on p. 23).

[34] Benjamin Bird, Milos Miljković, Stan Remiszewski, et al. “Infrared spectral histopathology (SHP): a novel diagnostic tool for the accurate classification of lung cancer”. *Laboratory investigation* 92.9 (2012), p. 1358 (cit. on p. 23).

[35] Shoji Kaminaka, Hiroya Yamazaki, Toshiaki Ito, Eiichi Kohda, and Hiro-o Hamaguchi. “Near-infrared Raman spectroscopy of human lung tissues: possibility of molecular-level cancer diagnosis”. *Journal of Raman Spectroscopy* 32.2 (2001), pp. 139–141 (cit. on pp. 24, 55, 70, 77, 80).

[36] Zhiwei Huang, Annette McWilliams, Harvey Lui, et al. “Near-infrared Raman spectroscopy for optical diagnosis of lung cancer”. *International journal of cancer* 107.6 (2003), pp. 1047–1052 (cit. on pp. 24, 55, 70–72, 77, 79, 110, 112).

[37] Michael A Short, Stephen Lam, Annette McWilliams, et al. “Development and preliminary results of an endoscopic Raman probe for potential in vivo diagnosis of lung cancers”. *Optics letters* 33.7 (2008), pp. 711–713 (cit. on p. 24).

[38] Nicholas D Magee, Julien S Villaumie, Eric T Marple, et al. “Ex vivo diagnosis of lung cancer using a Raman miniprobe”. *The Journal of Physical Chemistry B* 113.23 (2009), pp. 8137–8141 (cit. on pp. 24, 70, 71).

[39] Salvador González and Zeina Tannous. “Real-time, in vivo confocal reflectance microscopy of basal cell carcinoma”. *Journal of the American Academy of Dermatology* 47.6 (2002), pp. 869–874 (cit. on p. 25).

[40] Vinh Q Chung, Peter J Dwyer, Kishwer S Nehal, et al. “Use of ex vivo confocal scanning laser microscopy during Mohs surgery for nonmelanoma skin cancers”. *Dermatologic surgery* 30.12p1 (2004), pp. 1470–1478 (cit. on p. 25).

[41] Lorenzo Brancaleon, Anthony J Durkin, John H Tu, et al. “In vivo Fluorescence Spectroscopy of Nonmelanoma Skin Cancer”. *Photochemistry and Photobiology* 73.2 (2001), pp. 178–183 (cit. on p. 25).

[42] Masoud Panjehpour, Clark E Julius, Mary N Phan, Tuan Vo-Dinh, and Suzanne Overholt. “Laser-induced fluorescence spectroscopy for in vivo diagnosis of non-melanoma skin cancers”. *Lasers in Surgery and Medicine: The Official Journal of the American Society for Laser Medicine and Surgery* 31.5 (2002), pp. 367–373 (cit. on p. 25).

[43] John Paoli, Maria Smedh, Ann-Marie Wennberg, and Marica B Ericson. “Multiphoton laser scanning microscopy on non-melanoma skin cancer: morphologic features for future non-invasive diagnostics”. *Journal of Investigative Dermatology* 128.5 (2008), pp. 1248–1255 (cit. on p. 26).

[44] Bo Stenquist, Marica B Ericson, C Strandeberg, et al. “Bispectral fluorescence imaging of aggressive basal cell carcinoma combined with histopathological mapping: a preliminary study indicating a possible adjunct to Mohs micrographic surgery”. *British Journal of Dermatology* 154.2 (2006), pp. 305–309 (cit. on p. 26).

[45] Ann-Marie Wennberg, Fredrik Gudmundson, Bo Stenquist, et al. “In vivo detection of basal cell carcinoma using imaging spectroscopy.” *Acta dermato-venereologica* 79.1 (1999) (cit. on p. 26).

[46] Monika Gniadecka, HC Wulf, N Nymark Mortensen, O Faurskov Nielsen, and Daniel H Christensen. “Diagnosis of basal cell carcinoma by Raman spectroscopy”. *Journal of Raman Spectroscopy* 28.2-3 (1997), pp. 125–129 (cit. on p. 26).

[47] Patrick TT Wong, Sanford M Goldstein, Roy C Grekin, et al. “Distinct infrared spectroscopic patterns of human basal cell carcinoma of the skin”. *Cancer Research* 53.4 (1993), pp. 762–765 (cit. on p. 26).

[48] M Kazanci, P Roschger, EP Paschalis, K Klaushofer, and P Fratzl. “Bone osteonal tissues by Raman spectral mapping: orientation–composition”. *Journal of structural biology* 156.3 (2006), pp. 489–496 (cit. on p. 26).

[49] Mekhala Raghavan, Nadder D Sahar, Robert H Wilson, et al. “Quantitative polarized Raman spectroscopy in highly turbid bone tissue”. *Journal of biomedical optics* 15.3 (2010), p. 037001 (cit. on p. 26).

[50] Sonja Gamsjaeger, A Masic, P Roschger, et al. “Cortical bone composition and orientation as a function of animal and tissue age in mice by Raman spectroscopy”. *Bone* 47.2 (2010), pp. 392–399 (cit. on p. 26).

[51] Natalie Sheng Jie Lim, Zaribafzadeh Hamed, Chen Hua Yeow, Casey Chan, and Zhiwei Huang. “Early detection of biomolecular changes in disrupted porcine cartilage using polarized Raman spectroscopy”. *Journal of biomedical optics* 16.1 (2011), p. 017003 (cit. on p. 26).

[52] Admir Masic, Luca Bertinetti, Roman Schuetz, et al. “Observations of multiscale, stress-induced changes of collagen orientation in tendon by polarized Raman spectroscopy”. *Biomacromolecules* 12.11 (2011), pp. 3989–3996 (cit. on pp. 26, 109).

[53] Marek Janko, Polina Davydovskaya, Michael Bauer, Albert Zink, and Robert W Stark. “Anisotropic Raman scattering in collagen bundles”. *Optics letters* 35.16 (2010), pp. 2765–2767 (cit. on pp. 26, 101).

[54] Leonardo Galvis, John WC Dunlop, Georg Duda, Peter Fratzl, and Admir Masic. “Polarized Raman anisotropic response of collagen in tendon: towards 3D orientation mapping of collagen in tissues”. *PloS one* 8.5 (2013), e63518 (cit. on p. 26).

[55] Alex C-T Ko, Mark Hewko, Michael G Sowa, Cecilia CS Dong, Blaine Cleghorn, et al. “Detection of early dental caries using polarized Raman spectroscopy”. *Optics express* 14.1 (2006), pp. 203–215 (cit. on pp. 26, 121).

[56] Marie-Eve Rousseau, Thierry Lefevre, Lilyane Beaulieu, Tetsuo Asakura, and Michel Pézolet. “Study of protein conformation and orientation in silkworm and spider silk fibers using Raman microspectroscopy”. *Biomacromolecules* 5.6 (2004), pp. 2247–2257 (cit. on p. 26).

[57] George J Thomas Jr, James M Benevides, Stacy A Overman, et al. “Polarized Raman spectra of oriented fibers of A DNA and B DNA: anisotropic and isotropic local Raman tensors of base and backbone vibrations.” *Biophysical journal* 68.3 (1995), p. 1073 (cit. on p. 26).

[58] H Abramczyk, B Brozek-Pluska, and M Kopec. “Polarized Raman microscopy imaging: Capabilities and challenges for cancer research”. *Journal of Molecular Liquids* 259 (2018), pp. 102–111 (cit. on p. 26).

[59] Amuthachelvi Daniel, Aruna Prakasarao, Koteeswaran Dornadula, and Singaravelu Ganesan. “Polarized Raman spectroscopy unravels the biomolecular structural changes in cervical cancer”. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* 152 (2016), pp. 58–63 (cit. on p. 27).

[60] Elodie Ly, Anne Durlach, Frank Antonicelli, et al. “Probing tumor and peritumoral tissues in superficial and nodular basal cell carcinoma using polarized Raman microspectroscopy”. *Experimental dermatology* 19.1 (2010), pp. 68–73 (cit. on pp. 27, 85, 95, 98, 110).

[61] Elodie Ly Olivier Piot, Anne Durlach, Philippe Bernard, and Michel Manfait. “Polarized Raman microspectroscopy can reveal structural changes of peritumoral dermis in basal cell carcinoma”. *Applied spectroscopy* 62.10 (2008), pp. 1088–1094 (cit. on p. 27).

[62] ASTM E1840-96(2014). *Standard Guide for Raman Shift Standards for Spectrometer Calibration*. 2014 (cit. on p. 34).

[63] Narahara Chari Dingari, Gary L Horowitz, Jeon Woong Kang, Ramachandra R Dasari, and Ishan Barman. “Raman spectroscopy provides a powerful diagnostic tool for accurate determination of albumin glycation”. *PLoS One* 7.2 (2012), e32406 (cit. on p. 38).

[64] Neil J Everall. “Modeling and measuring the effect of refraction on the depth resolution of confocal Raman microscopy”. *Applied Spectroscopy* 54.6 (2000), pp. 773–782 (cit. on p. 38).

[65] J Trägårdh, K Macrae, C Travis, et al. “A simple but precise method for quantitative measurement of the quality of the laser focus in a scanning optical microscope”. *Journal of microscopy* 259.1 (2015), pp. 66–73 (cit. on p. 45).

[66] Takehiko Shimanouchi. “Tables of molecular vibrational frequencies part 5”. *Journal of Physical and Chemical Reference Data* 1.1 (1972), pp. 189–216 (cit. on pp. 46, 47).

[67] CA Murray and SB Dierker. “Use of an unintensified charge-coupled device detector for low-light-level Raman spectroscopy”. *JOSA A* 3.12 (1986), pp. 2151–2159 (cit. on p. 51).

[68] Jianhua Zhao, Harvey Lui, David I McLean, and Haishan Zeng. “Automated autofluorescence background subtraction algorithm for biomedical Raman spectroscopy”. *Applied spectroscopy* 61.11 (2007), pp. 1225–1232 (cit. on pp. 51, 59).

[69] Chad A Lieber and Anita Mahadevan-Jansen. “Automated method for subtraction of fluorescence from biological Raman spectra”. *Applied spectroscopy* 57.11 (2003), pp. 1363–1367 (cit. on pp. 51, 59, 77).

[70] Ranganathan Parthasarathy, Ganesh Thiagarajan, Xiaomei Yao, et al. “Application of multivariate spectral analyses in micro-Raman imaging to unveil structural/chemical features of the adhesive/dentin interface”. *Journal of biomedical optics* 13.1 (2008), p. 014020 (cit. on p. 51).

[71] Cláudia M Rocha, António S Barros, Ana M Gil, et al. “Metabolic profiling of human lung cancer tissue by 1H high resolution magic angle spinning (HRMAS) NMR spectroscopy”. *Journal of proteome research* 9.1 (2009), pp. 319–332 (cit. on pp. 55, 75).

[72] Nirmala Ramanujam. “Fluorescence spectroscopy of neoplastic and non-neoplastic tissues”. *Neoplasia* 2.1-2 (2000), pp. 89–117 (cit. on p. 55).

[73] Jaclyn Hung, Stephen Lam, Jean C Leriche, and Branko Palcic. “Autofluorescence of normal and malignant bronchial tissue”. *Lasers in surgery and medicine* 11.2 (1991), pp. 99–105 (cit. on pp. 55, 110).

[74] AC Croce and G Bottiroli. “Autofluorescence spectroscopy and imaging: a tool for biomedical research and diagnosis”. *European journal of histochemistry: EJH* 58.4 (2014) (cit. on p. 55).

[75] Kazuyuki Yano, Susumu Ohoshima, Yoshihiko Shimizu, Takeshi Moriguchi, and Hiroo Katayama. “Evaluation of glycogen level in human lung carcinoma tissues by an infrared spectroscopic method”. *Cancer letters* 110.1-2 (1996), pp. 29–34 (cit. on p. 55).

[76] Paul D Lewis, Keir E Lewis, Robin Ghosal, et al. “Evaluation of FTIR spectroscopy as a diagnostic tool for lung cancer using sputum”. *BMC cancer* 10.1 (2010), p. 640 (cit. on p. 55).

[77] Young-Kun Min, Tatsuya Yamamoto, Ehiichi Kohda, Toshiaki Ito, and Hiro-o Hamaguchi. “1064 nm near-infrared multichannel Raman spectroscopy of fresh human lung tissues”. *Journal of Raman Spectroscopy: An International Journal for Original Work in all Aspects of Raman Spectroscopy, Including Higher Order Processes, and also Brillouin and Rayleigh Scattering* 36.1 (2005), pp. 73–76 (cit. on pp. 55, 111).

[78] Dustin W Shipp, Faris Sinjab, and Ioan Notingher. “Raman spectroscopy: techniques and applications in the life sciences”. *Advances in Optics and Photonics* 9.2 (2017), pp. 315–428 (cit. on p. 55).

[79] Abigail S Haka, Karen E Shafer-Peltier, Maryann Fitzmaurice, et al. “Diagnosing breast cancer by using Raman spectroscopy”. *Proceedings of the National Academy of Sciences* 102.35 (2005), pp. 12371–12376 (cit. on p. 55).

[80] Tom C Bakker Schut, Peter J Caspers, Gerwin J Puppels, et al. “Discriminating basal cell carcinoma from its surrounding tissue by Raman spectroscopy”. *Journal of Investigative Dermatology* 119.1 (2002), pp. 64–69 (cit. on pp. 55, 84, 90).

[81] K Venkatakrishna, Jacob Kurien, Keerthilatha M Pai, et al. “Optical pathology of oral tissue: a Raman spectroscopy diagnostic method”. *CURRENT SCIENCE-BANGALORE-* 80.5 (2001), pp. 665–668 (cit. on p. 55).

[82] P Crow, A Molckovsky, N Stone, et al. “Assessment of fiberoptic near-infrared Raman spectroscopy for diagnosis of bladder and prostate cancer”. *Urology* 65.6 (2005), pp. 1126–1130 (cit. on p. 55).

[83] Senada Koljenovic, Tom C Bakker Schut, Jan P van Meerbeeck, et al. “Raman microspectroscopic mapping studies of human bronchial tissue”. *Journal of biomedical optics* 9.6 (2004), pp. 1187–1198 (cit. on p. 55).

[84] Ian R Lewis and Howell Edwards. *Handbook of Raman spectroscopy: from the research laboratory to the process line*. CRC Press, 2001 (cit. on p. 59).

[85] Chi Zhang, Delong Zhang, and Ji-Xin Cheng. “Coherent Raman scattering microscopy in biology and medicine”. *Annual review of biomedical engineering* 17 (2015), pp. 415–445 (cit. on p. 61).

[86] Nai-Teng Yu, John FR Kuck, and Carl C Askren. “Laser Raman spectroscopy of the lens in situ, measured in an anesthetized rabbit”. *Current eye research* 1.10 (1981), pp. 615–618 (cit. on p. 61).

[87] Sarah Marshall and John B Cooper. “Quantitative Raman Spectroscopy when the Signal-to-Noise is Below the Limit of Quantitation due to Fluorescence Interference: Advantages of a Moving Window Sequentially Shifted Excitation Approach”. *Applied spectroscopy* 70.9 (2016), pp. 1489–1501 (cit. on p. 61).

[88] Bernd-M Bussian and Wolfgang Härdle. “Robust smoothing applied to white noise and single outlier contaminated Raman spectra”. *Applied spectroscopy* 38.3 (1984), pp. 309–313 (cit. on p. 62).

[89] Mary Jo J Kerns, Marc A Darst, Thomas G Olsen, et al. “Shrinkage of cutaneous specimens: formalin or other factors involved?” *Journal of cutaneous pathology* 35.12 (2008), pp. 1093–1096 (cit. on p. 64).

[90] Po-Kuei Hsu, Hsu-Chih Huang, Chih-Cheng Hsieh, et al. “Effect of formalin fixation on tumor size determination in stage I non-small cell lung cancer”. *The Annals of thoracic surgery* 84.6 (2007), pp. 1825–1829 (cit. on p. 64).

[91] Thu Tran, Chandru P Sundaram, Clinton D Bahler, et al. “Correcting the shrinkage effects of formalin fixation and tissue processing for renal tumors: toward standardization of pathological reporting of tumor size”. *Journal of Cancer* 6.8 (2015), p. 759 (cit. on p. 65).

[92] Zanyar Movasaghi, Shazza Rehman, and Ihtesham U Rehman. “Raman spectroscopy of biological tissues”. *Applied Spectroscopy Reviews* 42.5 (2007), pp. 493–541 (cit. on pp. 71, 75, 79, 89, 92, 112).

[93] Ioan Notingher, Chris Green, Chris Dyer, et al. “Discrimination between ricin and sulphur mustard toxicity in vitro using Raman spectroscopy”. *Journal of the Royal Society Interface* 1.1 (2004), pp. 79–90 (cit. on p. 73).

[94] Nicholas Stone, Catherine Kendall, Jenny Smith, Paul Crow, and Hugh Barr. “Raman spectroscopy for identification of epithelial cancers”. *Faraday discussions* 126 (2004), pp. 141–157 (cit. on p. 73).

[95] Brooke D Beier and Andrew J Berger. “Method for automated background subtraction from Raman spectra containing known contaminants”. *Analyst* 134.6 (2009), pp. 1198–1202 (cit. on p. 77).

[96] Zhiwei Huang, ANNETTE McWILLIAMS, Stephen Lam, et al. “Effect of formalin fixation on the near-infrared Raman spectroscopy of normal and cancerous human bronchial tissues”. *International journal of oncology* 23.3 (2003), pp. 649–655 (cit. on p. 79).

[97] V Masilamani, K Al-Zhrani, M Al-Salhi, A Al-Diab, and M Al-Ageily. “Cancer diagnosis by autofluorescence of blood components”. *Journal of Luminescence* 109.3-4 (2004), pp. 143–154 (cit. on p. 79).

[98] Mohammad Al-Salhi, Vadivel Masilamani, Trinka Vijmasi, Hicham Al-Nachawati, and AP VijayaRaghavan. “Lung cancer detection by native fluorescence spectra of body fluids—a preliminary study”. *Journal of fluorescence* 21.2 (2011), pp. 637–645 (cit. on p. 80).

[99] Jarich W Spliethoff, Daniel J Evers, Houke M Klomp, et al. “Improved identification of peripheral lung tumors by using diffuse reflectance and fluorescence spectroscopy”. *Lung cancer* 80.2 (2013), pp. 165–171 (cit. on p. 80).

[100] Anna Cleta Croce, Gerardo Santamaria, Uliana De Simone, et al. “Naturally-occurring porphyrins in a spontaneous-tumour bearing mouse model”. *Photochemical & Photobiological Sciences* 10.7 (2011), pp. 1189–1195 (cit. on p. 80).

[101] GC Tang, A Pradhan, and RR Alfano. “Spectroscopic differences between human cancer and normal lung and breast tissues”. *Lasers in Surgery and Medicine* 9.3 (1989), pp. 290–295 (cit. on p. 80).

[102] Claudio Migliaresi and Antonella Motta. *Scaffolds for tissue engineering: Biological design, materials, and fabrication*. Pan Stanford, 2014 (cit. on p. 80).

[103] Patricia RM Rocco, Elnara M Negri, Pedro M Kurtz, et al. “Lung tissue mechanics and extracellular matrix remodeling in acute lung injury”. *American journal of respiratory and critical care medicine* 164.6 (2001), pp. 1067–1071 (cit. on p. 80).

[104] Mengyan Wang, Feng Long, Feng Tang, et al. “Autofluorescence Imaging and Spectroscopy of Human Lung Cancer”. *Applied Sciences* 7.1 (2016), p. 32 (cit. on p. 82).

[105] Marieke Pudlas, Steffen Koch, Carsten Bolwien, et al. “Raman spectroscopy: a noninvasive analysis tool for the discrimination of human skin cells”. *Tissue Engineering Part C: Methods* 17.10 (2011), pp. 1027–1040 (cit. on p. 84).

[106] Chad A Lieber, Shovan K Majumder, Darrel L Ellis, D Dean Billheimer, and Anita Mahadevan-Jansen. “In vivo nonmelanoma skin cancer diagnosis using Raman microspectroscopy”. *Lasers in Surgery and Medicine: The Official Journal of the American Society for Laser Medicine and Surgery* 40.7 (2008), pp. 461–467 (cit. on p. 84).

[107] Brian G Saar, L Rodrigo Contreras-Rojas, X Sunney Xie, and Richard H Guy. “Imaging drug delivery to skin with stimulated Raman scattering microscopy”. *Molecular pharmaceutics* 8.3 (2011), pp. 969–975 (cit. on p. 84).

[108] Chad A Lieber, Shovan K Majumder, D Dean Billheimer, Darrel L Ellis, and Anita Mahadevan-Jansen. “Raman microspectroscopy for skin cancer detection in vitro”. *Journal of biomedical optics* 13.2 (2009), p. 024013 (cit. on p. 84).

[109] M Larraona-Puy, A Ghita, A Zoladek, et al. “Discrimination between basal cell carcinoma and hair follicles in skin tissue sections by Raman microspectroscopy”. *Journal of Molecular Structure* 993.1-3 (2011), pp. 57–61 (cit. on pp. 85, 119).

[110] Kenny Kong, Christopher J Rowlands, Sandeep Varma, et al. “Diagnosis of tumors during tissue-conserving surgery with integrated autofluorescence and Raman scattering microscopy”. *Proceedings of the National Academy of Sciences* 110.38 (2013), pp. 15189–15194 (cit. on p. 85).

[111] Yongjian Zhu, Chun-Sik Choe, Sebastian Ahlberg, et al. “Penetration of silver nanoparticles into porcine skin ex vivo using fluorescence lifetime imaging microscopy, Raman microscopy, and surface-enhanced Raman scattering microscopy”. *Journal of biomedical optics* 20.5 (2014), p. 051006 (cit. on p. 85).

[112] Anthony T Tu. *Raman spectroscopy in biology. Principles and Applications 1*. Wiley, 1982 (cit. on p. 85).

[113] Elodie Ly, Olivier Piot, Anne Durlach, Philippe Bernard, and Michel Manfait. “Polarized Raman microspectroscopy can reveal structural changes of peritumoral dermis in basal cell carcinoma”. *Applied spectroscopy* 62.10 (2008), pp. 1088–1094 (cit. on p. 85).

[114] Michael A Short, Harvey Lui, David I McLean, et al. “Changes in nuclei and peritumoral collagen within nodular basal cell carcinomas via confocal micro-Raman spectroscopy”. *Journal of Biomedical Optics* 11.3 (2006), p. 034004 (cit. on pp. 90, 95, 98).

[115] Eugene A Bauer, James M Gordon, Max E Reddick, and Arthur Z Eisen. “Quantitation and immunocytochemical localization of human skin collagenase in basal cell carcinoma”. *Journal of Investigative Dermatology* 69.4 (1977), pp. 363–367 (cit. on p. 98).

[116] Franck Bonnier and HJ Byrne. “Understanding the molecular information contained in principal component analysis of vibrational spectra of biological systems”. *Analyst* 137.2 (2012), pp. 322–332 (cit. on p. 105).

[117] Valery Viktorovich Tuchin and V Tuchin. “Tissue optics: light scattering methods and instruments for medical diagnosis” (2007) (cit. on p. 109).

[118] Steven L Jacques, Jessica R Roman, and Ken Lee. “Imaging superficial tissues with polarized light”. *Lasers in Surgery and Medicine: The Official Journal of the American Society for Laser Medicine and Surgery* 26.2 (2000), pp. 119–129 (cit. on pp. 109, 114, 115).

[119] Zachary J Smith and Andrew J Berger. “Surface-sensitive polarized Raman spectroscopy of biological tissue”. *Optics letters* 30.11 (2005), pp. 1363–1365 (cit. on pp. 109, 115).

[120] Jürgen Popp, Valery V Tuchin, Arthur Chiou, and Stefan H Heinemann. *Handbook of biophotonics: Vol. 2: Photonics for health care*. Vol. 2. John Wiley & Sons, 2011 (cit. on p. 109).

[121] PA Mosier-Boss, SH Lieberman, and R Newbery. “Fluorescence rejection in Raman spectroscopy by shifted-spectra, edge detection, and FFT filtering techniques”. *Applied Spectroscopy* 49.5 (1995), pp. 630–638 (cit. on p. 110).

[122] Naiyan Huang, Michael Short, Jianhua Zhao, et al. “Full range characterization of the Raman spectra of organs in a murine model”. *Optics express* 19.23 (2011), pp. 22892–22909 (cit. on p. 110).

[123] Roger L Mcmullen, Guojin Zhang, and Timothy Gillece. “Quantifying hair shape and hair damage induced during reshaping of hair”. *J Cosmet Sci* 66 (2015), pp. 379–409 (cit. on p. 110).

[124] Valery P Zakharov, Ivan A Bratchenko, Dmitry N Artemyev, et al. “Lung neoplasm diagnostics using Raman spectroscopy and autofluorescence analysis”. *Journal of Biomedical Photonics & Engineering* 1.1 (2015) (cit. on p. 110).

[125] Chetan A Patil, Isaac J Pence, Chad A Lieber, and Anita Mahadevan-Jansen. “1064 nm dispersive Raman spectroscopy of tissues with strong near-infrared autofluorescence”. *Optics letters* 39.2 (2014), pp. 303–306 (cit. on p. 111).

[126] Kun Chen, Hongyuan Zhang, Haoyun Wei, and Yan Li. “Improved Savitzky–Golay-method-based fluorescence subtraction algorithm for rapid recovery of Raman spectra”. *Applied optics* 53.24 (2014), pp. 5559–5569 (cit. on p. 111).

[127] Pavel Matousek and Nicholas Stone. “Emerging concepts in deep Raman spectroscopy of biological tissue”. *Analyst* 134.6 (2009), pp. 1058–1066 (cit. on p. 111).

[128] Chad A Lieber and Mustafa H Kabeer. “Characterization of pediatric Wilms’ tumor using Raman and fluorescence spectroscopies”. *Journal of pediatric surgery* 45.3 (2010), pp. 549–554 (cit. on p. 111).

[129] Vanitha Sankaran, Matthew J Everett, Duncan J Maitland, and Joseph T Walsh. “Comparison of polarized-light propagation in biological tissue and phantoms”. *Optics Letters* 24.15 (1999), pp. 1044–1046 (cit. on p. 114).

[130] Ryan CN Studinski and I Alex Vitkin. “Methodology for examining polarized light interactions with tissues and tissue-like media in the exact backscattering direction”. *Journal of Biomedical Optics* 5.3 (2000), pp. 330–338 (cit. on p. 115).

[131] Gilbert Jarry, Elisa Steimer, Vivien Damaschini, et al. “Coherence and polarization of light propagating through scattering media and biological tissues”. *Applied optics* 37.31 (1998), pp. 7357–7367 (cit. on p. 115).

[132] Zhiwei Huang, Seng Khoon Teh, Wei Zheng, et al. “Integrated Raman spectroscopy and trimodal wide-field imaging techniques for real-time in vivo tissue Raman measurements at endoscopy”. *Optics letters* 34.6 (2009), pp. 758–760 (cit. on p. 120).

[133] JCC Day, R Bennett, B Smith, et al. “A miniature confocal Raman probe for endoscopic use”. *Physics in Medicine & Biology* 54.23 (2009), p. 7077 (cit. on p. 120).

[134] Lisa J McCawley and Lynn M Matrisian. “Matrix metalloproteinases: multifunctional contributors to tumor progression”. *Molecular medicine today* 6.4 (2000), pp. 149–156 (cit. on p. 121).