

CRANFIELD UNIVERSITY



OLIVIA JACKSON

VIBRATIONAL SPECTROSCOPY FOR THE RAPID AND EARLY DIAGNOSIS OF
LEUKAEMIAS AND LYMPHOMAS

CRANFIELD HEALTH

PhD THESIS

Academic Year: 2013-14

Academic Supervisors: Professor Nicholas Stone & Dr Conrad Bessant
Clinical Supervisors: Dr Adam Rye, Dr Richard Lush & Dr Keith McCarthy

November 2013

CRANFIELD UNIVERSITY

CRANFIELD HEALTH

PhD THESIS

Academic Year: 2013-14

OLIVIA JACKSON

VIBRATIONAL SPECTROSCOPY FOR THE RAPID AND EARLY DIAGNOSIS OF
LEUKAEMIAS AND LYMPHOMAS

Academic Supervisors: Professor Nicholas Stone & Dr Conrad Bessant
Clinical Supervisors: Adam Rye, Richard Lush & Keith McCarthy

November 2013

This thesis is submitted in partial fulfilment of the requirements for the degree of
Doctor of Philosophy

© Cranfield University, 2013. All rights reserved. No part of this publication may
be reproduced without the written permission of the copyright holder.

ABSTRACT

This thesis aimed to investigate vibrational spectroscopies for the identification of biochemical markers of leukaemias and lymphomas. In a preliminary study using the blood proteins albumin, fibrinogen and globulin, Drop Coating Deposition Raman Spectroscopy was explored and extended for use with Fourier Transform infrared spectroscopy for leukaemia blood sample analysis. Due to low sample volumes and minimal preparation required it was identified as a potential alternative to blood centrifugation to obtain the buffy coat for analysis. These studies identified that it was capable of detecting low levels of protein from small, highly concentrated droplets. Thus this method, alongside cytopsin centrifugation, was used for the spectroscopic analysis of different blood fractions. Due to the low number of lymphoma samples obtained, only a feasibility study is outlined in this thesis.

Samples were collected from leukaemia patients and healthy volunteers. Infrared and Raman spectra were measured of whole blood and buffy coat samples cytopsin onto slides and whole blood and plasma pipetted by drop coating deposition. Multivariate statistical analysis was employed to extract key spectral differences between the pathologies and develop classification models for diagnosing chronic lymphoblastic leukaemia from previously treated and untreated patient groups. Principal component analysis followed by linear discriminant analysis was employed to identify the largest variances in the data and leave one sample out cross validation evaluated the performance of the spectral models measured on different blood components in diagnosing leukaemia. The buffy coat infrared model correctly classified 59% of the spectra, and blood droplet Raman 62%. The treated and untreated groups were then combined, which improved classification to 83% for buffy coat infrared and 71% for blood droplet Raman. These findings highlight the potential of drop coating deposition spectroscopy of whole blood for leukaemia diagnosis, although further work is required to achieve a clinically validated method.

ACKNOWLEDGEMENTS

I would like to thank Gloucestershire Hospitals NHS Foundation Trust (GHNHSFT) and Cranfield University for their collaboration. I would also like to thank the Leukaemia and Intensive Chemotherapy Charity (LINC) for funding this project.

My thanks also go to my supervisors, Nick Stone, Conrad Bessant, and Michael Cauchi, the Biophotonics Research Group, particularly Gavin Lloyd, Catherine Kendall and Allison Bates for their continual help and support and the haematology laboratory at Gloucestershire Royal Hospital for their time and allowing me to use their equipment.

Finally I wish to thank my family and friends for their support and encouragement throughout the project, particularly my husband who without his continual support and praise I would not have succeeded after such a difficult year.

Page Contents

iii	Abstract
iv	Acknowledgments
v	Contents
ix	Figures
xii	Tables
xiv	Notation
1	1. Introduction
1	1.1. Background
3	1.2. Haematological Malignancies
5	1.2.1. Development of Haematological Malignancies
9	1.2.2. Classification
12	1.2.3. Staging
14	1.2.4. Incidence
16	1.2.5. Aetiology
18	1.2.6. Current Diagnostic Methodology
21	1.2.7. Deficiencies
24	1.2.8. Clinical Need
25	1.3. Optical Diagnostic Methods
25	1.3.1. The Electromagnetic Spectrum
26	1.3.2. The Interaction of Light with Matter
28	1.3.3. Infrared Absorption Spectroscopy
29	1.3.4. Fluorescence Spectroscopy
30	1.3.5. Elastic Scattering Spectroscopy
31	1.3.6. Inelastic Scattering Spectroscopy (Raman)
32	1.3.7. Conclusion
33	1.4. Aims and Objectives
35	2. Raman and Fourier Transform Infrared Spectroscopy
35	2.1. Fourier Transform Infrared Spectroscopy

38	2.1.1. Biomedical and Haematological Applications of FTIR
41	2.2. Raman Spectroscopy
46	2.2.1. Biomedical and Haematological Applications of Raman
48	2.2.2. Drop Coating Deposition Spectroscopy
50	2.3. Data Analysis
50	2.3.1. Spectral Pre-Processing
51	2.3.2. Peak Ratio Analysis
51	2.3.3. Chemometric Analysis
53	2.4. Conclusion
55	2.5. Summary
56	3. Materials and Methods
57	3.1. Optimisation of Spectral Measurements: Identifying Parameters
58	3.1.1. Fourier Transform Infrared Spectroscopy
62	3.1.2. Raman Spectroscopy
66	3.1.3. Raman System Calibration
68	3.2. Evaluating Drop Coating Deposition Spectroscopy (DCDS) for Microanalysis of Protein Solutions
68	3.2.1. Sample Preparation
69	3.2.2. Raman, Infrared Transmittance and Infrared Reflectance Spectroscopy
69	3.2.3. Identifying the Position and Absorbance of the Amide I Peak
70	3.3. Determining the Optimal Ring Characteristics for DCDS Using a Series of Albumin Solutions
70	3.3.1. Sample Preparation
71	3.3.2. Infrared and Raman Spectral Line Maps
72	3.3.3 Amide I Peak Analysis
72	3.4. Comparing DCDS and Cytospin Centrifugation as Blood Sample Preparation Methods in the Diagnosis of Leukaemia Using FTIR and Raman Spectroscopy
72	3.4.1. Ethics
72	3.4.2. Sample Collection
73	3.4.3. Sample Preparation

73	3.4.4. Methods
73	3.4.4.1. Isolating Mononuclear Cells from Anti-Coagulated Whole Blood by Density Gradient Centrifugation
76	3.4.4.2. Drop Coating Deposition Spectroscopy
76	3.4.4.3. Spectroscopic Analysis
77	3.4.4.4. Data Analysis
78	3.5. Lymphoma Diagnosis Using FTIR and Raman Spectroscopy
78	3.5.1. Ethics
78	3.5.2. Sample Collection & Preparation
78	3.5.3. Methods
78	3.5.3.1. Cytospin Centrifugation
78	3.5.3.2. Drop Coating Deposition Spectroscopy
78	3.4.5.2. Spectroscopic Analysis
79	4. Results and Discussion: Blood Protein Analysis
79	4.1. Evaluating Drop Coating Deposition Spectroscopy (DCDS) for Microanalysis of Protein Solutions
79	4.1.1. Sample Preparation Using DCDS
83	4.1.2. Raman, Infrared Transmittance and Infrared Reflectance Spectroscopy
86	4.1.3. Identifying the Position and Absorbance of the Amide I Peak
94	4.1.4. Summary
95	4.2. Determining the Optimal Ring Characteristics for DCDS Using a Series of Albumin Solutions
95	4.2.1 Physical properties
97	4.2.2 Amide I Peak Analysis
99	4.2.3. Determining Minimum Ring Width
101	4.2.4. Summary
102	4.3. Discussion
104	5. Results and Discussion: Leukaemia Analysis
104	5.1 Screening Study
105	5.1.1. Morphological Analysis

105	5.1.2. Empirical Analysis
111	5.2. Mean Spectra
116	5.3. Difference Spectra
131	5.4. Principal Component Analysis
140	5.5. Linear Discriminant Analysis
143	5.6. Leave One Sample Out Cross Validation
146	5.7. Two Group Model
151	5.8. Discussion
152	6. Results and Discussion: Lymphoma Analysis
152	6.1. Physical Properties
154	6.2. Spectral Analysis of Lymphoma Samples
158	6.3. Discussion
160	7. Conclusions
160	7.1. DCDS Method
160	7.2. Leukaemia Classification Model
163	7.3. Lymphoma
163	7.4. Clinical Applications
164	7.5. Future Work
166	References
185	Appendix A. 3CCN, Cancer Patients: A Simplified Pathway
187	Appendix B. 1. Characteristic Infrared Peaks in the Fingerprint region 900-1800cm ⁻¹
201	Appendix B. 2. Characteristic Raman Shifts in the Fingerprint region 900- 1800cm ⁻¹
213	Appendix C. Correspondence from South West REC re ethical approval
215	Appendix D. Histopaque-1077 Product Information Sheet
217	Appendix E. Pure Protein Spectra
219	Appendix F. Mean Spectra
222	Appendix G. List of Publications, Presentations and Prizes

6	Figure 1.1. Blood Cell Lineages That Arise From A Pluripotent Stem Cell
8	Figure 1.2. The Lymphatic System
9	Figure 1.3. Genetic Regulation of Proliferation and Apoptosis
15	Figure 1.4. Five-Year Survival Rates for Leukaemia, Hodgkin Lymphoma and Non-Hodgkin Lymphoma in the UK, from 1975 to 2005
22	Figure 1.5. Process of Diagnostic Analysis of Haematological Malignancies
23	Figure 1.6. Three Counties Cancer Network Patient Pathway
26	Figure 1.7 The Electromagnetic Spectrum of Radiation
28	Figure 1.8 Energy Level Diagram of Mid-IR, Rayleigh, Stokes and Anti-Stokes Raman
37	Figure 2.1. Structural Layout of a Common FTIR Spectrometer
42	Figure 2.2. Structural Layout of a Common Raman Spectrometer
43	Figure 2.3. Raman and IR Active Molecules
49	Figure 2.4. Image of a Typical Droplet Deposited on a CaF ₂ Slide
59	Figure 3.1. Mean FTIR Spectra of Lymph Node Tissue Sections To Identify the Effect of Increasing Spectral Resolution on Quality of Spectra
60	Figure 3.2. Mean FTIR Spectra of Lymph Node Tissue Sections To Identify the Effect of Increasing the Number of Scans Per Pixel on Quality of Spectra
61	Figure 3.3. Mean FTIR Spectra of Lymph Node Tissue Sections To Identify the Effect of Increasing Aperture Size on Quality of Spectra
63	Figure 3.4. Mean Raman Spectra of Lymph Node Tissue Sections To Identify the Effect of Different Objectives on Quality of Spectra
64	Figure 3.5. Mean Raman Spectra of Lymph Node Tissue Sections To Identify the Effect of Spectral Acquisition Time on Quality of Spectra
65	Figure 3.6. Mean Raman Spectra of Lymph Node Tissue Sections To Identify the Effect of Increasing the Number of Accumulations on Quality of Spectra
67	Figure 3.7. Spectra of Raman Calibration Standards
70	Figure 3.8. Schematic Diagram Showing the Stages of Line Profile Analysis

74	Figure 3.9. Conical Tube 1
75	Figure 3.10. Conical Tube 2
76	Figure 3.11. The Cytospin Mounts
80	Figure 4.1. White Light Images of Albumin Droplets
81	Figure 4.2. White Light Images of Fibrinogen Droplets
82	Figure 4.3. White Light Images of γ -Globulin Droplets
85	Figure 4.4. Typical Spectra from Albumin Ring A1
87	Figure 4.5. Peak Absorption and Position Plots of the Amide I Peak for Albumin Droplet A2 Measured in IR Transmittance
88	Figure 4.6. Line Profile Plots of Albumin Rings Measured in IR Transmittance
90	Figure 4.7. 3D Spectral Plots of Albumin Rings A and A2 Measured in IR Transmittance
91	Figure 4.8. Line Profile Plots of Albumin Rings Measured in Raman
92	Figure 4.9. Line Profile Plot of Albumin Ring A Measured in IR Reflectance
96	Figure 4.10. White Light Image of Albumin Droplets on CaF ₂ Slide
96	Figure 4.11. The Effect of Ring Width as Albumin Concentration Increases Across All Three Techniques
98	Figure 4.12. Average Amide I Peak Position Across All Rings Measured in Raman, IR Transmittance & IR Reflectance
99	Figure 4.13. Line Profile Plot of Ring A7
100	Figure 4.14. Measurement of Ring Width by Calculating Mean and Standard Deviation of the Amide I Peak at Data Points Across the Ring Width
108	Figure 5.1. Histogram Plots of Amide I/ Amide II Ratios
109	Figure 5.2. Plasma IR Spectra
111	Figure 5.3. Mean Healthy IR Spectra of All Four Sample Types
112	Figure 5.4. Mean Healthy Raman Spectra of All Four Sample Types
114	Figure 5.5. Mean Buffy Coat IR Spectra of the Three Pathologies
115	Figure 5.6. Mean Buffy Coat Raman Spectra of the Three Pathologies
117	Figure 5.7. Difference Spectra for Buffy Coat IR
119	Figure 5.8. Difference Spectra for Blood Cytospin IR
121	Figure 5.9. Difference Spectra for Blood Droplet IR

123	Figure 5.10. Difference Spectra for Buffy Coat Raman
125	Figure 5.11. Difference Spectra for Blood Cytospin Raman
127	Figure 5.12. Difference Spectra for Blood Droplet Raman
129	Figure 5.13. Difference Spectra for Plasma Raman
133	Figure 5.14. Plots of the F-Ratio for Each PC
135	Figure 5.15. PC Loads for Each Dataset as Selected by ANOVA
136	Figure 5.16. PC Loads for IR Buffy Coat
139	Figure 5.17. PC Score Plots
142	Figure 5.18. LDA Plots
145	Figure 5.19. Percentage of Spectra In Each Dataset Correctly Classified by LOSOCV
148	Figure 5.20. Percentage of Spectra In Each Dataset In the Two Group Model Correctly Classified by LOSOCV
152	Figure 6.1. White Light Images of the First FNAB Received
154	Figure 6.2. Cytospin White Light Images
155	Figure 6.3. Mean Spectra of Sample 1
156	Figure 6.4. PBS Spectra
158	Figure 6.5. Mean Spectra of FNAB Samples

Page Tables

7	Table 1.1. Leucocyte Classification
10	Table 1.2. Classification of Haematological Malignancies
15	Table 1.3. UK Incidence and Mortality Rates per 100,000 people for 1975 and 2007
19	Table 1.4. Advantages and Disadvantages of Diagnostic Methods
44	Table 2.1. Comparison of FTIR and Raman in Cancer Diagnostics
58	Table 3.1. FTIR and Raman Parameter Options
58	Table 3.2. SNR for FTIR Spectra
62	Table 3.3. SNR for Raman Spectra
68	Table 3.4. Protein Concentrations in Human Blood
69	Table 3.5. Protein Dilutions
71	Table 3.6. Albumin Dilutions
93	Table 4.1. Amide I Peak Positions (cm^{-1})
105	Table 5.1. Samples Obtained
106	Table 5.2. Amide I/ Amide II Ratios
110	Table 5.3. Number of Spectra Removed From Each Sample Due to Saturation
115	Table 5.4. Raman Peak Assignments
118	Table 5.5. Assignment of Difference Spectra Peaks for Buffy Coat IR
120	Table 5.6. Assignment of Difference Spectra Peaks for Blood Cytospin IR
122	Table 5.7. Assignment of Difference Spectra Peaks for Blood Droplet IR
124	Table 5.8. Assignment of Difference Spectra Peaks for Buffy Coat Raman
126	Table 5.9. Assignment of Difference Spectra Peaks for Blood Cytospin Raman
128	Table 5.10. Assignment of Difference Spectra Peaks for Blood Droplet Raman
130	Table 5.11. Assignment of Difference Spectra Peaks for Plasma Raman
132	Table 5.12. F_{crit} Values of the Seven Sample Groups
143	Table 5.13. LOSOCV Results for Buffy Coat IR
143	Table 5.14. LOSOCV Results for Blood Cytospin IR

144	Table 5.15. LOSOCV Results for Blood Droplet IR
144	Table 5.16. LOSOCV Results for Buffy Coat Raman
144	Table 5.17. LOSOCV Results for Blood Cytospin Raman
144	Table 5.18. LOSOCV Results for Blood Droplet Raman
145	Table 5.19. LOSOCV Results for Plasma Raman
146	Table 5.20. Two Group LOSOCV Results for Buffy Coat IR
146	Table 5.21. Two Group LOSOCV Results for Blood Cytospin IR
147	Table 5.22. Two Group LOSOCV Results for Blood Droplet IR
147	Table 5.23. Two Group LOSOCV Results for Buffy Coat Raman
147	Table 5.24. Two Group LOSOCV Results for Blood Cytospin Raman
147	Table 5.25. Two Group LOSOCV Results for Blood Droplet Raman
148	Table 5.26. Two Group LOSOCV Results for Plasma Raman
149	Table 5.27. Overall Performance of Datasets in the Two and Three Group Models
150	Table 5.28. FTIR LOSOCV Results Using 12PCs
150	Table 5.29. Raman LOSOCV Results Using 12PCs
150	Table 5.30 Comparison of the Performance of the Complete and Conservative Two Group Models

NOTATION

α	Alpha
β	Beta
μl	Microlitre
μm	Micrometers
%	Percentage
3CCN	Three Counties Cancer Network
A	Absorbance
ALA	5-Aminolevulinic Acid
ALL	Acute Lymphoblastic Leukaemia
AML	Acute Myeloid Leukaemia
ANOVA	Analysis of Variance
AU	Arbitrary Units
BFU	Burst Forming Unit
CCD	Charge-Coupled Device
CFU	Colony Forming Unit
CLL	Chronic Lymphoblastic Leukaemia
cm	Centimetre
CML	Chronic Myeloid Leukaemia
CT	Computerised Tomography
DCDS	Drop Coating Deposition Spectroscopy
DCDRS	Drop Coating Deposition Raman Spectroscopy
DNA	Deoxyribonucleic Acid
EDTA	Ethylenediaminetetracetic Acid
EMS	Electromagnetic Spectrum
ESS	Elastic Scattering Spectroscopy
FAB	French American British
F_{crit}	Critical Value of F
FISH	Fluorescence <i>in situ</i> Hybridisation
FNAB	Fine Needle Aspirate Biopsy
FTIR	Fourier Transform Infrared

GHNHSFT	Gloucestershire Hospitals NHS Foundation Trust
GP	General Practitioner
H&E	Haematoxylin and Eosin
HIV	Human Immunodeficiency Virus
HL	Hodgkin Lymphoma
HPD	Hematoporphyrin Derivative
I	Incident Light
I_0	Light emitted from sample
ISS	Inelastic Scattering Spectroscopy
IR	Infrared
LDA	Linear Discriminant Analysis
LINC	Leukaemia and Intensive Chemotherapy Charity
LOOCV	Leave One Out Cross Validation
LOSOCV	Leave One Sample Out Cross Validation
LSS	Light Scattering Spectroscopy
MCL	Mast Cell Leukaemia
MDT	Multidisciplinary Team
MIR	Mid Infrared
mL	Millilitre
MZL	Marginal Zone Lymphoma
NA	Not applicable
NHL	Non-Hodgkin Lymphoma
NIR	Near Infrared
NK	Natural Killer cells
PBS	Phosphate Buffered Saline
PC	Principal Component
PCA	Principal Component Analysis
PCL	Plasma Cell Leukaemia
PCR	Polymerase Chain Reaction
PET	Positron Emission Tomography
PIS	Patient Information Sheet
PpIX	Protoporphryn IX
PTP	Previously Treated Patients

PUP	Previously Untreated Patients
REC	Research Ethics Committee
RNA	Ribonucleic Acid
SD	Standard Deviation
SN	Sensitivity
SP	Specificity
T	Transmittance
UK	United Kingdom
USAF	United States Air Force
UV	Ultraviolet
WHO	World Health Organisation

1. INTRODUCTION

The purpose of this thesis was to investigate the application of vibrational spectroscopy, Raman and Fourier Transform Infrared (FTIR), in the early diagnosis of Leukaemia and Lymphoma. The first two chapters illustrate the current literature surrounding haematological malignancies, along with their current diagnostic methods and the deficiencies associated with them. Also reviewed here is the literature relating to the theories, uses and potential application of vibrational spectroscopies in the early diagnosis of haematological malignancies and explores the methods for data analysis. Chapter 3 contains the experimental methods used. This includes the preliminary work into the techniques, which has been vital in producing protocols for the leukaemia and lymphoma studies. Chapters 4, 5 and 6 contain the results and a discussion of the work carried out and chapter 7 concludes the project with further ideas for the experiments and future work suggested.

1.1. Background

In England, there were nearly 270,000 newly diagnosed cases of cancer in 2010 – an increase of 1.52 % from 2009 (Office for National Statistics, 2012). It is the principal cause of death in the Western World and is responsible for over seven million deaths worldwide each year. Overall, cancer is diagnosed in one in three people; killing one in four thus the earlier a diagnosis is made the earlier treatment can commence, leading to a better prognosis for the patient (Cancer Research UK, 2012a).

Leukaemia and lymphoma combined account for approximately 20,000 incidences of cancer and 9,000 deaths in the United Kingdom (UK) each year (Cancer Research UK, 2012b & 2011a). Unfortunately these cancers are the most common to occur in children and although their incidence has not increased over the last forty years, they currently account for 6 % of all childhood deaths (Office for National Statistics, 2011a). Some of the current methods used to diagnose

these cancers identify molecular variations at the cellular level using morphological, immunological and molecular techniques. However, they can be inaccurate due to the complexity of the interpretation involved, which means either repetitive or multiple methods are required to be carried out by highly trained professionals.

Vibrational spectroscopies are analytical methods that have useful applications in the clinical field. They can be used to identify the biochemical composition of cells and tissues, which is useful in identifying the changes that occur at the molecular level within disease. The biochemical differences between 'normal' and a disease state can therefore be identified as molecular markers. These markers are disease-specific, as the information obtained in spectroscopy is unique and the spectra obtained are effectively a 'biochemical fingerprint'. By using the two vibrational spectroscopies Raman and FTIR in combination, highly detailed information can be identified as they have specificities to particular chemical bonds. These techniques are automated, reliable and effective at probing the underlying biochemical changes that occur. When used in collaboration with multivariate statistical analysis more specific and accurate information can be extracted from the data obtained leading to more precise and robust data models that represent the cancers at the biochemical level (Kendall *et al*, 2009). Thus, these techniques are explored in this thesis to provide a valuable method for diagnosis of haematological malignancies that can improve upon the methods currently used.

With the aid of the Leukaemia and Intensive Chemotherapy Charity (LINC), the Biophotonics Research Group within Gloucestershire Hospitals National Health Service Foundation Trust (GHNHSFT) and Cranfield University, this thesis aims to identify new, rapid and improved techniques for the early diagnosis of two types of haematological malignancy: leukaemia and lymphoma.

1.2. Haematological Malignancies

Leukaemias, lymphomas and myelomas are classified as haematological malignancies with leukaemia and lymphoma having a high incidence in children (Lightfoot & Roman, 2004, Mostaço-Guidolin *et al*, 2011 & Zelig *et al*, 2011). However, overall they affect a larger number of adults, with the most common forms of leukaemia and lymphoma most prevalent in people over the age of fifty (Cancer Research UK, 2011b, 2012b, 2013a & 2013b). The main difference between leukaemias and lymphomas is the way in which they progress. Leukaemias proliferate as single cells in the blood or lymph, whereas lymphomas do so as a solid tumour mass within lymphoid tissue such as bone marrow (Mostaço-Guidolin *et al*, 2011, & Hoffbrand *et al*, 2010). It has been questioned which stage of differentiation gives rise to leukaemia; however, it appears most likely that the type of leukaemia developed is dependent on the cell lineage and stage of differentiation at which the abnormality occurred (Celso *et al*, 2009). Myeloma, which is not the focus of this thesis, is cancer of the plasma cells found in bone marrow.

Leukaemia

Leukaemia is a disease that has been well recognised throughout history, with Bennett first describing it in the mid nineteenth century (Bennett, 1845 & Piller, 2001). It is a cancer of the leucocytes, found in the lymph nodes and spleen, and the bone marrow (Babrah *et al*, 2009 & Sheng *et al*, 2013). It develops through the proliferation of immature leucocytes, which are vital in the fight against infection. Thus if the leucocytes are abnormal then the risk of infection is higher which in turn becomes more difficult to fight (Sahu *et al*, 2006). A build up of leucocytes in parts of the lymphatic system and liver can cause abdominal swelling. Too many leucocytes also reduce the space for other types of blood cells such as erythrocytes (red blood cells) and megakaryocytes (platelets). A lack of erythrocytes gives rise to anaemia and not enough megakaryocytes leads to problems with excess bleeding (Zelig *et al*, 2011 & Hoffbrand *et al*, 2010).

There are several subtypes of leukaemia, which differ in the type of leucocytes affected and the time it takes to develop. The two main forms of leukaemia are

acute, which is aggressive proliferation and accumulation of immature leucocytes that develops quickly at any age, and *chronic* which is a slow proliferation of mature leucocytes that develops slowly in adults. These are further subdivided into myeloid and lymphoblastic (also referred to as lymphocytic) according to the cells that are affected (Jennings & Foon, 1997, Piller, 2001 & Hoffbrand *et al*, 2010). The most common type of leukaemia in the Western world is chronic lymphoblastic leukaemia (CLL), which affects B cells (Shultz *et al*, 1996 & Cancer Research UK, 2013a). It is probably the most studied form of leukaemia due to the fact that it is more accessible to researchers and that it can be studied over time.

Lymphoma

Lymphoma is a cancer of the lymphatic system, commonly affecting the spleen, bone marrow and peripheral lymph nodes (Ronson *et al*, 2006). It develops through the proliferation of lymphocytes, which accumulate in lymph nodes to form a tumour. There are two main types, Hodgkin Lymphoma (HL), named after Thomas Hodgkin, and non-Hodgkin Lymphoma (NHL), with NHL being the most prevalent (Ansell & Armitage, 2005, Cancer Research UK 2011b & 2013b). The major difference between HL and NHL can be identified under a microscope as the cells affected in HL have a particular binucleate appearance, these are known as Reed Sternberg cells (Jennings & Foon, 1997). These cells are not present in NHL.

Both forms of lymphoma affect cells of the lymph nodes, which are found all over the body and are connected via lymph vessels. Unfortunately this network can often lead to the spread of lymphoma not just to other lymph nodes but also to the blood and organs (Isabelle *et al*, 2008). These cancers often present with enlarged lymph nodes, owing to the accumulation of leucocytes within the lymphatic system. The nodes that they affect can help to differentiate between HL and NHL. HL tends to affect the lymph nodes in the neck first whereas NHL more commonly affects lymph nodes in the tonsils or thymus. However, this is not always true and cannot be relied upon as a method of classification (Hoffbrand *et al*, 2010).

1.2.1. Development of Haematological Malignancies

All blood cells develop from a haematopoietic pluripotent stem cell, which differentiates into cells of the myeloid or lymphoid progenitor (Figure 1.1). Stem cells have one of three fates: they can remain inactive, undergo cell division and differentiation, or undergo apoptosis (Kwong & Chan, 1988). This process is very tightly controlled and regulated by external signals such as transcription factors, cytokines and interleukins with glycoprotein hormones acting as growth factors in the differentiation and maturation of the cells (Hoffbrand *et al*, 2010). If the genes responsible for these processes do not function properly, a stem cell may undergo the wrong fate. For example, the genes responsible for apoptosis may be 'faulty' permitting an abnormal cell to continue dividing and differentiating, giving rise to cancer (Attar & Scadden, 2004). Cancer may also arise if immature blood cells do not receive the required signals to mature, thus generating a surplus of these cells in the body and a lack of the vital, functional mature cells.

The process of blood cell development is haemopoiesis. In a fetus, haemopoiesis originates in the yolk sac, progressing to the liver and spleen, then to the bone marrow before a baby is born. During infancy the bone marrow remains the site of blood cell development. However, during adult life when marrow becomes more fatty, this process tends to occur throughout the central skeleton in such places as the ribs, pelvis and femur. Approximately one trillion blood cells are produced each day, of which at least ten billion are vital in the maintenance of homeostasis. During any stage of differentiation these cells can become abnormal giving rise to a haematological malignancy, cancers that affect the blood, bone marrow and lymph nodes. (Cantor & Orkin, 2001, Attar & Scadden, 2004 & Hoffbrand *et al*, 2010)

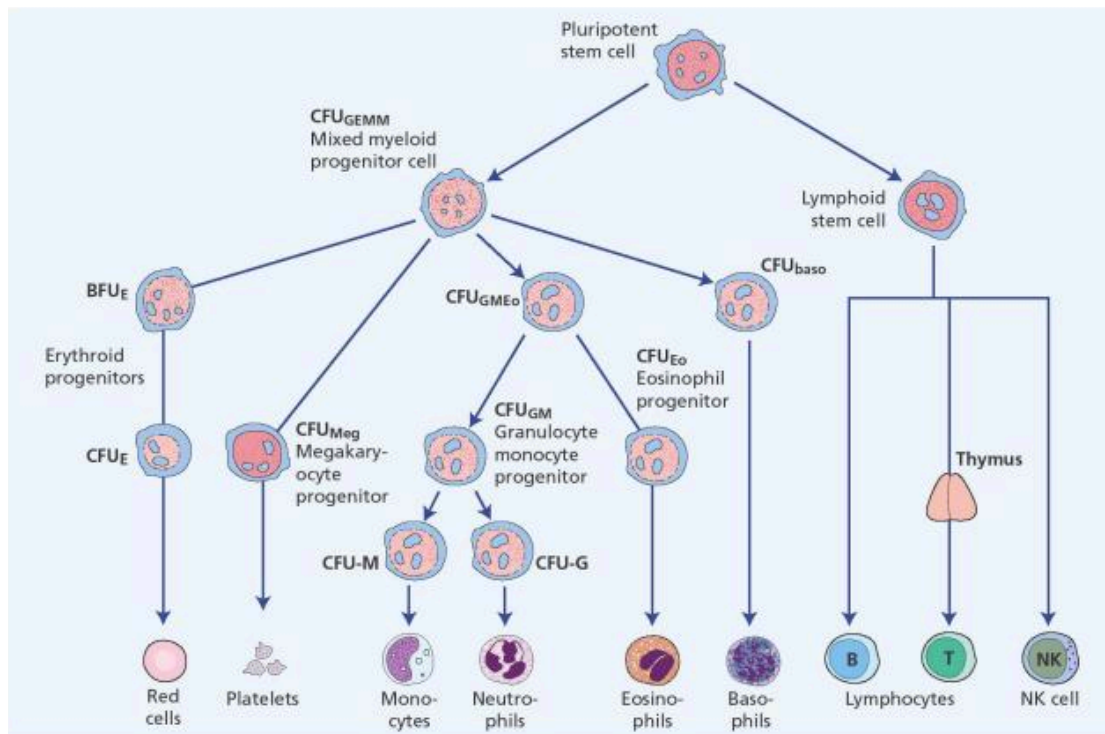


Figure 1.1. Blood Cell Lineages That Arise From A Pluripotent Stem Cell
 (Source: Hoffbrand *et al*, 2010)

Once blood cells have matured, they are released into the bloodstream where they are able to carry out their designated functions. Erythrocytes (red blood cells) circulate oxygen around the body whilst megakaryocytes (platelets) are involved in blood coagulation. Leucocytes (white blood cells) are the largest of the blood cells, contributing to less than 1 % of the total blood volume (Waugh & Grant, 2005). They are cells of the immune system that aid the body to fight against infection, and can be divided into two major groups: phagocytes and lymphocytes. These can be further subdivided as shown in Table 1.1. The phagocytes, as their name suggest, are responsible for phagocytosis, a process in which these cells engulf and digest foreign bodies such as bacteria. The lymphocytes also target foreign bodies via the production of antibodies (Hoffbrand *et al*, 2010). Haematological malignancies occur through abnormalities in leucocyte development, and the type of leukaemia or lymphoma that arises is dependent on the cell type that is affected.

Table 1.1. Leucocyte Classification

Phagocytes	Lymphocytes
<ul style="list-style-type: none">• Granulocytes<ul style="list-style-type: none">- Neutrophils- Eosinophils- Basophils• Monocytes	<ul style="list-style-type: none">• B Cells• T Cells• Natural Killer (NK) Cells

The lymphatic system is a large circulatory network found in the body consisting of the bone marrow, lymph, lymph nodes, lymph vessels and lymph organs (Figure 1.2). It is responsible for the production and transportation of lymphocytes to sites of infection. All leucocytes circulate in the blood through capillaries and can pass into lymph organs via lymph vessels (Waugh & Grant, 2005). Owing to the largely connective nature of the lymphatic system and its vast expansion throughout the body, haematological malignancies have the potential to spread rapidly to many organs thus emphasising the importance and the urgency in diagnosing these cancers early (Isabelle *et al*, 2008).

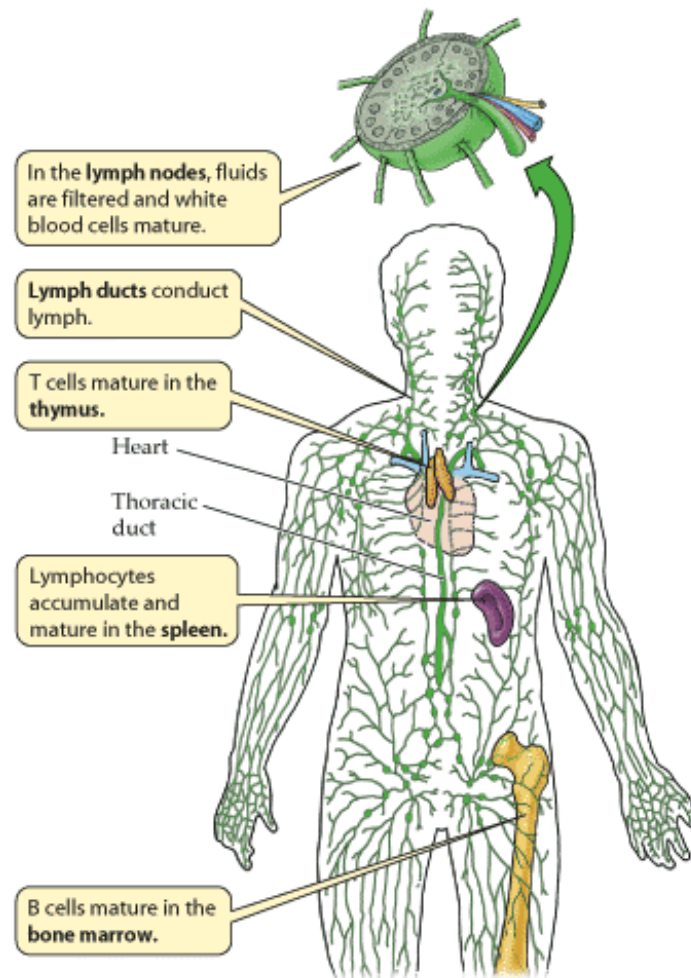


Figure 1.2. The Lymphatic System; the vast lymphatic network is represented in green, the major organs involved and their functions are highlighted (Source: Purves *et al*, 2004)

Haematopoietic cells turn over rapidly in the body so as to maintain homeostasis (Debatin *et al*, 2003). If these cells fail to respond to the external stimuli responsible for their growth, an excess accumulation of immature, non-functional cells can arise. It is vital that these cells are removed and thus apoptotic pathways are necessary to destroy them. Haematological malignancies tend to result when there are problems with apoptosis as there is an increased number of immature cells which are unable to carry out normal cellular functions. Failures in the apoptotic pathways can create genetically unstable environments whereby genetic mutations can arise that support tumour cell

survival. Proto-oncogenes, which regulate growth and proliferation, and tumour-suppressor genes, which control apoptosis of malignant cells, can be activated or inactivated by a cell in order to maintain the required balance of genes for normal cell growth (Figure 1.3). An unbalance of these genes results in tumour cell growth (Hoffbrand *et al*, 2010).

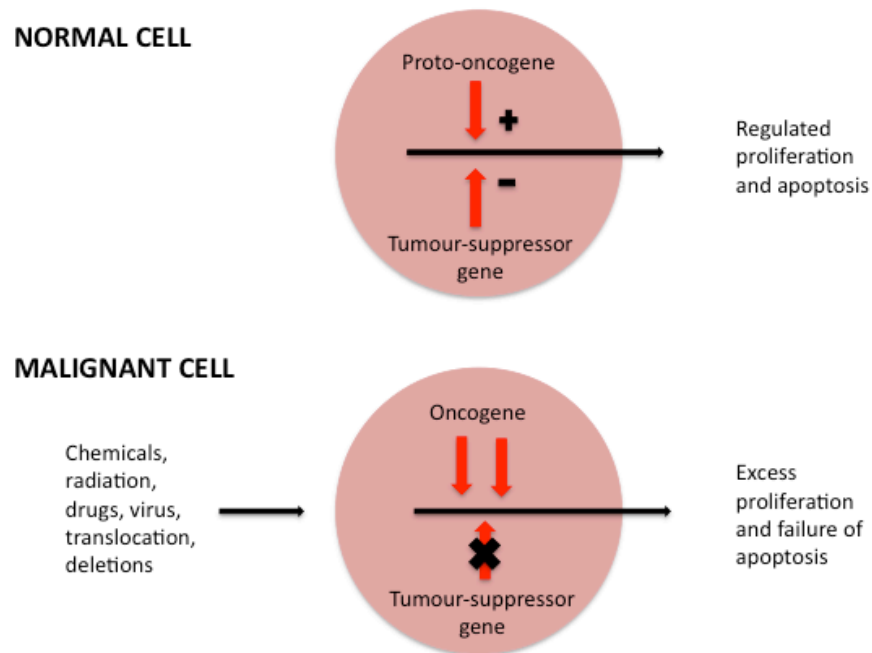


Figure 1.3. Genetic Regulation of Proliferation and Apoptosis; normal cell proliferation depends on balance between proto-oncogenes and tumour-suppressor genes, which is unstable in malignant cells (adapted from Hoffbrand *et al*, 2010)

1.2.2. Classification

The distinctions between each type of cancer are based on morphology, and clinical presentation combined with molecular and immunophenotyping methods (Ottensmeier, 2001). Leukaemias are broadly divided into myeloid or lymphoblastic and these can either be acute or chronic according to how aggressive they are. High grade, or acute, leukaemias are very aggressive and

require chemotherapy in order to be treated. Low grade, or chronic, cancers are indolent, meaning that they are less active. This type of cancer tends to progress slowly but is less curable and treatment is normally required to reduce accompanying symptoms (Andrus & Strickland, 1998).

Lymphomas are generally referred to as either Hodgkin or non-Hodgkin although there are several subtypes. NHL can be classified as B-cell or T-cell diseases according to the cell-type involved. Again they are classified as either chronic or acute according to how aggressive they are (Andrus & Strickland, 1998). The major subtypes of these cancers, along with information regarding cells affected, onset and age of risk are summarised in Table 1.2.

Table 1.2. Classification of Haematological Malignancies

Haematological Malignancy	Leucocyte involved	Onset	Risk Age
Leukaemia - <i>Acute myeloid</i>	Myeloid –monocyte/ granulocyte	Days –weeks	>65
- <i>Acute lymphoblastic</i>	B –Lymphocyte	Days – weeks	<20
- <i>Chronic myeloid</i>	Myeloid – granulocyte	Months – years	40 – 60
- <i>Chronic lymphoblastic</i>	B –Lymphocyte	Months – years	>60
Hodgkin lymphoma	B –Lymphocyte (Reed – Sternberg)	Depends on type	15-29 & >60
Non-Hodgkin lymphoma	B - & T – Lymphocytes	Depends on type	50 – 60

Generated from UK cancer registries statistics (Cancer Research UK, 2011b, 2013a & 2013b)

Many of the leukaemias and lymphomas in Table 1.2 can be subdivided or classified into even smaller groups. Different countries use different classification systems.

Leukaemia Classification

Two systems are used to classify acute myeloid leukaemia (AML), the French-American-British (FAB) and the World Health Organisation (WHO). The FAB system classifies AML according to the morphological appearance of the cells under a microscope along with any chromosomal changes and the antibody markers on the cell. AML is classified into eight different groups, from M₀, M₁... to M₇, with M₀ having the best prognosis:

- M₀, M₁, M₂ – Myeloblastic Leukaemia (50 % of all cases)
- M₃ – Promyelocytic Leukaemia (10 % of cases)
- M₄ – Myelomonocytic Leukaemia (20 % of cases)
- M₅ – Acute Monocytic Leukaemia (15 % of cases)
- M₆ – Acute Erythroleukaemia (rare)
- M₇ – Acute Megakaryocytic Leukaemia (rare)

The WHO system is slightly different. It looks at the type of myeloid cell that is abnormal as well as any genetic changes, whether more than one type of blood cell is abnormal, whether the AML developed following a blood disorder and finally whether AML developed after other cancer treatment.

Acute Lymphoblastic Leukaemia (ALL) used to be classified into three groups, L₁, L₂, L₃, according to the FAB system. This system identifies changes to cell size, nuclear size and shape, number of nucleoli and amount of cytoplasm. Now it is classified as follows according to the WHO system.

- Pre (precursor) B Cell - most common (70 % of cases)
- Mature B Cell or Burkitt Type
- Pre (Precursor) T Cell – more common in men and young adults

Classification of chronic leukaemias is much simpler. The most common forms are chronic myeloid leukaemia (CML) and chronic lymphocytic leukaemia (CLL).

More rare forms include prolymphocytic leukaemia and hairy cell leukaemia. (Lilleyman *et al*, 1986, Jennings & Foon, 1997, Woolf, 1998 & Szczepański *et al*, 2003).

Lymphoma Classification

Hodgkin lymphoma (HL) is classified into classical types or nodular lymphocyte predominant type according to the WHO system. This system takes into account the cellular appearance under a microscope. There are four classical types of Hodgkin Lymphoma, characterised by the presence of Reed-Sternberg cells:

- Nodular Sclerosing (60 % of cases)
- Mixed Cellularity (15 % of cases)
- Lymphocyte Rich (10 % of cases)
- Lymphocyte Depleted (very rare)

Nodular lymphocyte predominant type accounts for 20 % of cases. They differ from the classical type in that they have very few Reed-Sternberg cells but instead have 'popcorn' shaped cells (Cancer Research UK, 2013c).

Non-Hodgkin Lymphoma (NHL) is classified according to the WHO system. There are over sixty different subtypes and thus classification is difficult. However, generally they are classified according to whether they are high grade or low grade, B-cell or T-cell, their cellular appearance (large, small, grouped or diffuse), what proteins or markers are present on the cell surface and if there are any genetic abnormalities (Cancer Research UK, 2013d).

1.2.3. Staging

Haematological malignancies can also be staged according to their progression. This is vital in providing the most suitable treatment to the patient.

Leukaemia Staging

There is no standard staging system for the acute leukaemias. Patients are broadly staged as:

- Untreated – newly diagnosed patients with an abnormal full blood count, >5 % of cells in bone marrow are blasts and signs and symptoms of leukaemia.
- Remission – these patients have a normal full blood count, <5 % blasts in their bone marrow and no signs or symptoms of leukaemia.

There are three stages of CML: chronic, accelerated and blast phases (Cancer Research UK, 2013a):

- *Chronic phase* – leukaemia is most stable and still developing slowly. Ten per cent of the blood cells in bone marrow will be immature blasts. About 90 % of people are in the chronic phase when they are diagnosed.
- *Accelerated phase* – symptoms such as tiredness, weight loss and painful stomach due to an enlarged spleen can arise. Between 10 and 30 % of the blood cells will be immature blasts.
- *Blast phase* – Rapid and aggressive leukaemia, also called acute phase or blast crisis. Over 30 % of the blood cells will be immature blasts and the disease will have spread to other organs.

There are two systems used to stage CLL, one is the Rai system used in America and the other is known as the Binet System, which is used in the UK and Europe. The Binet System categorises the cancer into three stages, A, B and C (Cancer Research UK, 2013a):

- *Stage A* – fewer than three groups of enlarged lymph nodes and a high white blood cell count. Patient is unlikely to show symptoms.
- *Stage B* – more than three groups of enlarged lymph nodes are present and there is a high white blood cell count. Patient will be tired and poorly or show no symptoms.
- *Stage C* – patient will have enlarged lymph nodes or spleen; high white blood cell count; low red blood cell and platelet counts. Low levels of red blood cells (anaemia), blood clots, infections, weight loss and night sweats are common.

Lymphoma Staging

The system used to stage both types of lymphoma in the UK and Europe is the Ann Arbor system, which divides the cancer into four stages, 1, 2, 3 and 4 and within these four stages cancers can also be further divided into type A or B. Type A refers to an absence of symptoms whereas B refers to a patient suffering with symptoms (Ansell & Armitage, 2005, Cancer Research UK, 2011b & 2013b):

- *Stage 1* – only a single tumour resides and has not spread. One group of lymph nodes is affected or lymphoma occurs in just one organ of the body.
- *Stage 2* – two or more groups of lymph nodes are affected or an organ and one or more groups of lymph nodes but these are on the same side of the diaphragm, i.e. the tumour has not spread far.
- *Stage 3* – lymph nodes are affected on both sides of the diaphragm.
- *Stage 4* – lymphoma has spread outside the lymph nodes and reached organs such as the liver, bone or lungs.

1.2.4. Incidence

The occurrence of haematological malignancies has changed significantly over the last thirty years (Table 1.3). Unfortunately they are increasing in both incidence and mortality, along with many other cancers (Ronson *et al*, 2006). The prevalence of these cancers varies amongst age groups; Hodgkin lymphoma is most common in people under thirty with a high five-year survival rate of 80 % compared to only 50 % for non-Hodgkin. HL has a bimodal distribution with the highest incidences being around the mid twenties and in the over sixties. The incidence of NHL increases with age and is more common in males than females. Overall, leukaemia is the thirteenth most common cancer in the UK with over 7,000 people diagnosed in 2007. Nine in ten of those diagnosed were adults and only 420 of them children. Chronic myeloid leukaemia (CML) and acute lymphoblastic leukaemia (ALL) account for one in ten of all leukaemia's, with ALL being the most common in children. Fortunately survival rates have increased since the 1970's, although this could be improved significantly if the cancers were diagnosed sooner (Figure 1.4).

Table 1.3. UK Incidence and Mortality Rates per 100,000 people for 1975 and 2007

	Leukaemia		Hodgkin Lymphoma		Non-Hodgkin Lymphoma	
	Incidence	Mortality	Incidence	Mortality	Incidence	Mortality
1975	7.4	5.9	2.8	1.4	5.5	3.3
2007	9.5	4.9	2.6	0.4	14.2	5.2

Generated from UK cancer registries statistics (Cancer Research UK, 2011a, 2012b, & Office for National Statistics, 2011b)

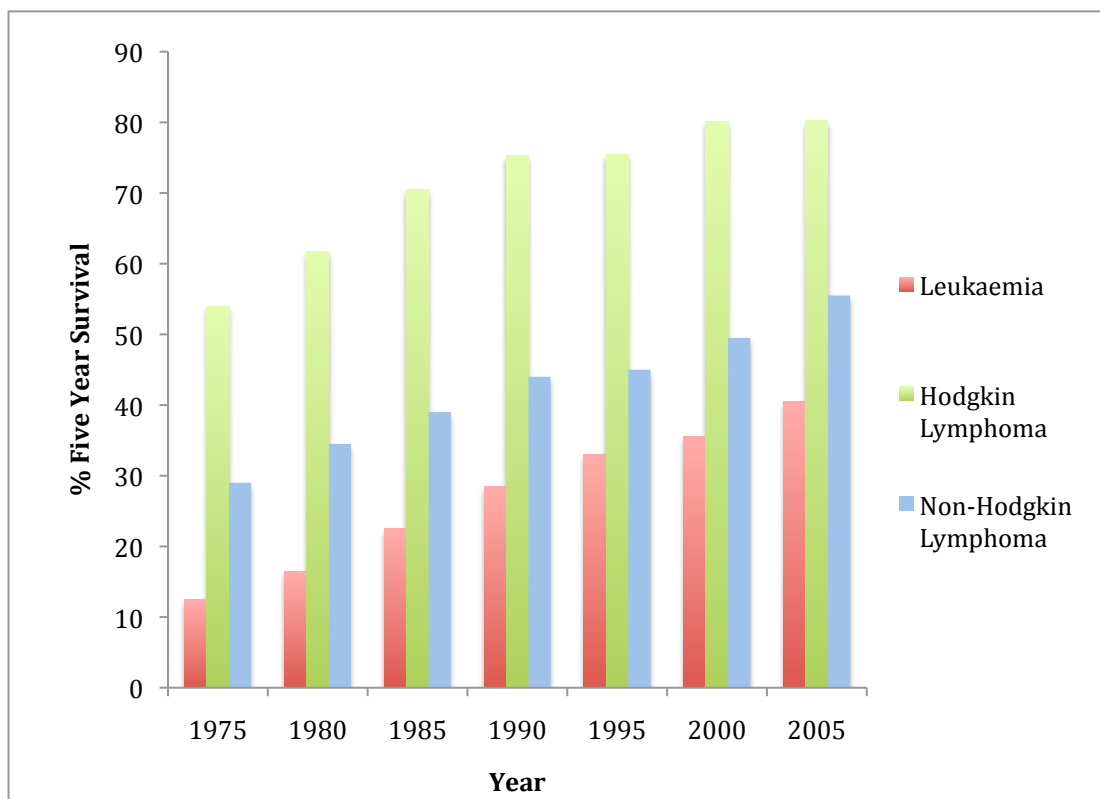


Figure 1.4. Five-Year Survival Rates for Leukaemia, Hodgkin Lymphoma and Non-Hodgkin Lymphoma in the UK, from 1975 to 2005 (Source: Cancer Research UK, 2010a, 2010b, 2010c & Office For National Statistics, 2011b)

Comparing the data in Table 1.3 with Figure 1.4 it is clear to see that in 2007, there were 7.4 cases of leukaemia per 100,000 people, an increase of 28 % from 1975. The mortality rates for the disease reduced by 17 % which although is an improvement, when compared to the five-year survival rate of 40 % of cases, it is still poor. Both the incidence and mortality rates for Hodgkin lymphoma have improved over the last thirty years, alongside the five-year survival rate. In 1975, 53.9 % of sufferers survived for five years, better than the survival rate for leukaemia in 2005. Today over 80 % of Hodgkin Lymphoma patients survive for more than five years. The incidence of non-Hodgkin Lymphoma has almost tripled since 1975, which is a significant growth in the disease. The mortality also increased by 58 %, again a significant increase in the last thirty years. Fortunately the five-year chance of survival has improved from 29 to 55.5 % of cases, although like leukaemia this is still low compared to other cancers. These statistics show that leukaemia and non-Hodgkin lymphoma still have a fair way to go to be on par with the survival rates of diseases such as testicular cancer (97.1 %) and breast cancer (84.2 %), which means more research is needed into the aetiology, potential molecular markers for diagnosis as well as improved methods for diagnosis and treatment (Office for National Statistics, 2011b). Trends similar to these can be seen across the world, mainly in the developed countries where diagnosis and treatment is available. In countries where medical help is less readily available the incidence appears the same, however mortality is higher and the one, five, and ten-year survival rates are significantly lower.

1.2.5. Aetiology

The aetiology of these cancers remains unknown as only a small number of haematological malignancies arise from an inherited mutation (Segel & Lichtman, 2004). However, the risk factors for a small proportion of subjects have been identified (Vineis *et al*, 1996). These can be of genetic or environmental influence and can affect both children and adults. Risk factors can range from exposure to materials such as benzene, pesticides, and past chemotherapy, which are mostly thought to have immunosuppressive qualities,

to smoking, drinking, being overweight and even using particular hair dyes (Vineis *et al*, 1996, Woolf, 1998 & Lightfoot & Roman, 2004). The majority of children that are diagnosed with a haematological malignancy would most likely not have been exposed to risk factors such as these so early in life. It may be that the parents have been exposed to such risks and Lightfoot & Roman (2004) suggest that exposure to either parent may create germ cell mutations in their children.

It has been suggested by Smith *et al* (2009) that the risk of leukaemia occurring in childhood is higher in girls and in those that were overweight at birth, and their work supports the theory that leukaemia originates prenatally. However, a previous report by Lightfoot & Roman (2004) stated that haematological malignancies are more commonly seen in boys than girls, with the majority of lymphomas occurring in boys. It has also been reported that there is a higher risk of NHL developing in people with low immunity, such as human immunodeficiency virus (HIV) sufferers (Vineis *et al*, 1996). According to Vineis *et al* (1996) the highest incidence rates for lymphoma occur in developed areas whilst the lowest are concerned with underdeveloped areas. This phenomenon may be due to the risk factors associated such as exposure to pesticides, radiation and benzene, which are more commonly used in the developed world. However, it may just be due to the lack of statistics available, or that less people are diagnosed in underdeveloped areas due to a deficiency in the medical care available.

Although numerous studies have been carried out into these risk factors, there is currently not enough substantial evidence to be able to identify who is at a possible risk of developing a haematological malignancy. Those risk groups that have been identified vary considerably amongst researchers. Risks, such as exposure to benzene do not necessarily mean that a haematological malignancy will develop and having an overweight daughter at birth does not predetermine that they will develop leukaemia. One may not have been exposed to any of the risk factors mentioned and still develop a haematological malignancy or one may have been exposed to several factors and never develop cancer. Many papers

actually contradict each other on their findings, rendering it difficult to identify particular groups of people who are at high risk of developing cancer (Lightfoot & Roman, 2004).

1.2.6. Current Diagnostic Methodology

In the UK there are currently no screening programmes for haematological malignancies thus more novel diagnostic techniques are crucial for good prognosis (Ansell & Armitage, 2005). Current “gold standard” techniques used to diagnose haematological malignancies rely on the identification of morphological features, such as cell size, nuclear size, and the presence of multiple nuclei (Dukor, 2002). Whilst the morphology may aid in the identification of a malignancy, it is not always sufficient in classifying the subtypes and stages of a disease and thus is used in combination with other methods (Dukor, 2002, Richards & Jack, 2003 & Szczepański *et al*, 2003). Techniques often used include immunohistochemistry, cytogenetics, electron microscopy, and biochemical markers (Ottensmeier, 2001 & Dukor, 2002). Morphology combined with immunophenotyping is generally sufficient in the majority of patients and further methods may only be required in the diagnosis of more difficult cases.

In order to diagnose leukaemias it is common for Haematoxylin and Eosin (H&E) stained blood samples or bone marrow aspirations to be examined under a microscope, thus providing information on the cellular morphology. Immunophenotyping is commonly carried out following morphological identification in order to gain a more definitive diagnosis. Diagnosing Hodgkin and non-Hodgkin lymphoma is very similar. They too rely on the morphological analysis of H&E stained samples, but excised lymph node tissue is commonly used. This method is quite invasive, as minor surgery is carried out in order to gain access to lymph node tissue (Sahu & Mordechai, 2005). More recently fine needle aspirate biopsies (FNAB) have been introduced, which reduces how invasive and uncomfortable diagnosis is to the patient. Other molecular methods used in conjunction with morphology and immunophenotyping methods are highlighted in Table 1.4.

Table 1.4. Advantages and Disadvantages of Diagnostic Methods

Diagnostic Method	Use/ Advantage	Disadvantage
<i>Clinical Features</i>	Identifies cancer according to symptoms, e.g. enlarged lymph nodes.	Not necessarily indicative of cancer thus may lead to unnecessary diagnostic analysis. Can cause excessive costs, time and stress.
<i>Morphological Analysis</i>	Identifies abnormalities in H&E stained samples.	Subjective, costly and time-consuming. Not able to classify or grade cancers.
<i>Lymph Node Biopsy</i>	Histological identification of enlarged lymph nodes and tissues.	Invasive, uncomfortable and can generate false positives.
<i>Bone Marrow Biopsy</i>	Identifies spread of cancer to bone marrow.	Invasive, uncomfortable and can generate false positives.
<i>Cytogenetics</i>	Identifies specific genetic abnormalities.	Invasive and requires fresh cells.
<i>Immunophenotyping</i>	Uses antibodies to identify presence of cell specific proteins related to cancer. Can specifically identify stages of cancer.	Can only be used for cancers with known immunophenotype, commonly used alongside cytogenetics to identify chromosomal differences.

<i>Molecular Methods (PCR/ FISH/ Southern Blotting)</i>	Identify specific subtypes of cancer utilising their gene expression.	Requires fresh tissue as deoxyribonucleic acid (DNA) degradation can occur, time consuming, labour intensive and can generate false positives.
<i>Positron Emission Tomography (PET)</i>	Uses radiolabelled glucose to identify presence of disease following treatment.	Expensive.
<i>Computerised Tomography (CT) Scan</i>	Used in conjunction with PET, identifies spread of cancer to other organs.	Exposure to ionising radiation.
<i>Vibrational Spectroscopies (FTIR/ Raman)</i>	Identify changes in biochemical composition, non-destructive, quick, reliable, easily interpretable.	FTIR limited by water so tissue must be extracted and dried – makes <i>in vivo</i> difficult (this can be overcome by using in combination with Raman). Raman signal weak, fluorescence can interfere and sample degradation may occur.

PCR: Polymerase Chain Reaction, FISH: Fluorescent *in situ* Hybridisation (Sources: Alberts *et al*, 2002, Richards & Jack, 2003, Szczepański *et al*, 2003, Ansell & Armitage, 2005, Griffiths & de Haset, 2007, Smith & Dent, 2008, Kendall *et al*, 2009 & Hoffbrand *et al*, 2010)

1.2.7. Deficiencies

The current “gold standard” techniques may not be sufficient in the diagnosis of all haematological malignancies as the morphological analysis undertaken by histopathologists are subjective, varying from one doctor to the next. According to Szczepański *et al*, (2003) between 5-10 % of leukaemia cases require further analysis in order to give a definitive result and of these, 25 % remain unresolved following further molecular testing. Sample interpretation can be complex if an inadequate amount of malignant cells are present. More samples may then be required from the patient, which can be uncomfortable, disconcerting and time consuming. The combination and number of methods that may be used increases the time and cost of the diagnostic procedures.

Richards and Jack (2003) highlight the requirement of a multidisciplinary team in obtaining a diagnosis, stating that for the process to run smoothly all professionals involved are best placed in close proximity to each other within a hospital. Unfortunately this is not always possible; particularly within GHNHSFT where there are two main hospital sites located over eight miles apart. Figure 1.5 indicates the complex process of taking a sample and preparing it for analysis, followed by the different molecular diagnostic routes that can be taken.

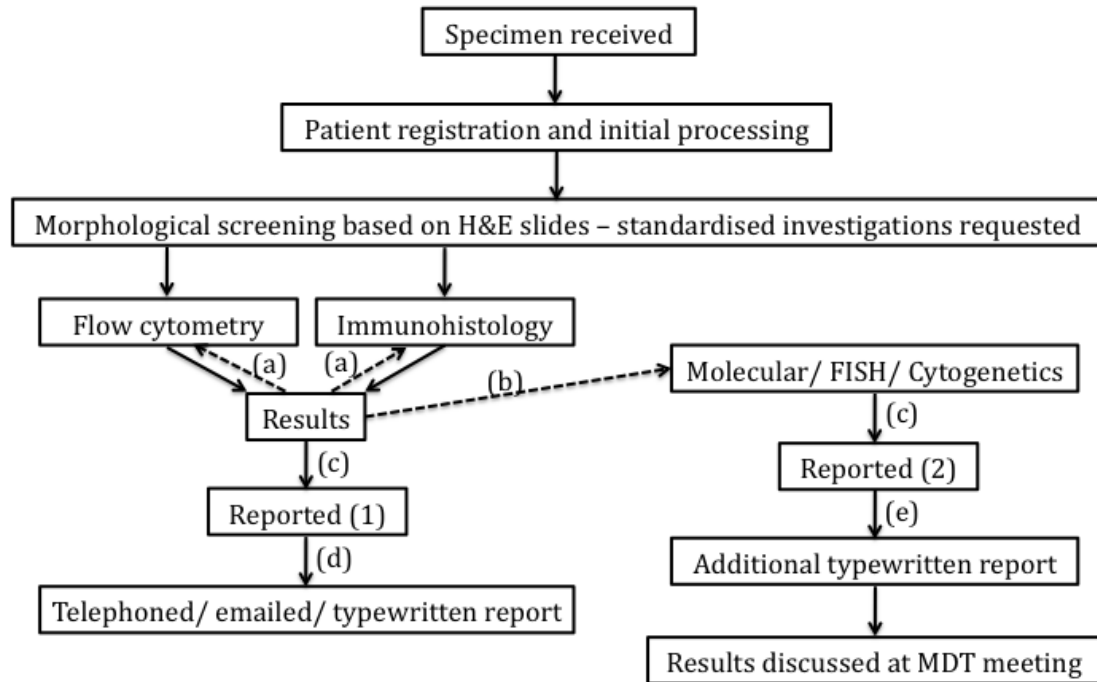


Figure 1.5. Process of Diagnostic Analysis of Haematological Malignancies; (a) repeated flow cytometry or immunohistology required, (b) initial results obtained may suggest further molecular investigations, (c) results checked by senior member of team, (d) haematological malignancy diagnoses are checked by a further member of team to ensure correct diagnosis, (e) another report produced to highlight additional information regarding the malignancy (adapted from Richards & Jack, 2003)

Within GHNHSFT a document has been drawn up for patients to indicate the timescale from hospital referral and initial diagnosis to cancer treatment and after care (Figure 1.6). Referrals are generally made by a GP (General Practitioner) to the hospital, which can take up to 14 days (stage 1). Following this, the initial services are provided. This involves diagnostic tests ranging from blood tests to biopsies and X-rays or other scans. The type and number of tests carried out determines the length of time taken and can significantly alter the cost (stage 2). The diagnosis and treatment plan is then reviewed at a multidisciplinary team (MDT) meeting, involving all of the medical professionals that will work with the patient (stage 3), before it is discussed with the patient. It is at this point that a maximum of 31 days is allowed before the first treatment

must be provided. The type of treatment can range from surgery to chemotherapy or radiotherapy (stage 4). The overall time allowed for an urgent case, such as an acute form of leukaemia, from referral to the first treatment is 62 days, or two months, which can be longer if non-urgent. The patient pathway then continues with follow-up appointments, and support, which can be provided in a variety of ways (stage 5) (Three Counties Cancer Network, 2005).

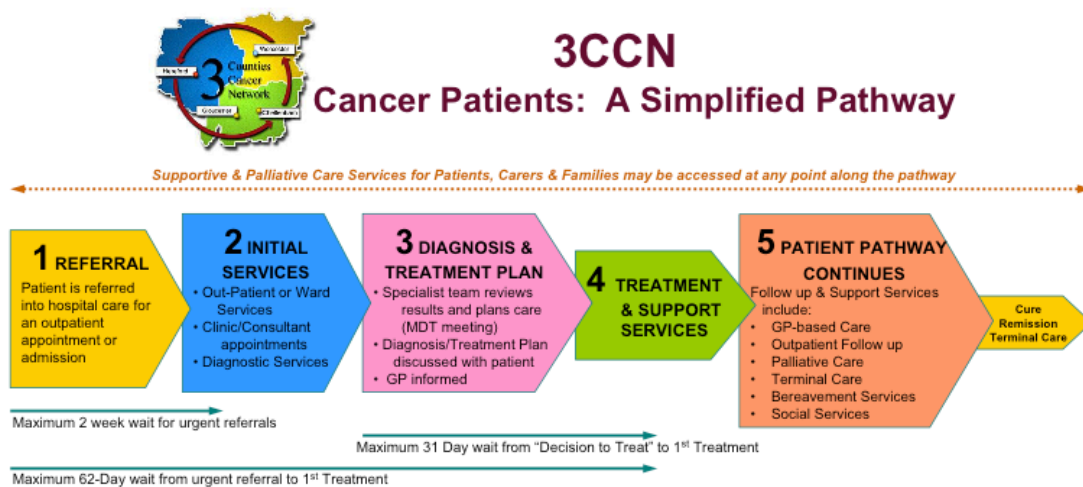


Figure 1.6. Three Counties Cancer Network Patient Pathway (3CCN). Full document can be found in Appendix A. (Source: Three Counties Cancer Network, 2005)

Not only can the process of diagnosis be time consuming and costly, the methods used can also be invasive. Excising tissue samples requires minor surgery and even extracting blood can be traumatic for some patients (Sahu & Mordechai, 2005). FNAB is a very good diagnostic tool for both HL and high grade NHL with a success rate of 79-90 %. Although it is less invasive, it appears to be less successful in diagnosing low grade NHL (Kocjan, 2005). According to Hehn *et al* (2004) the use of FNABs to diagnose lymphomas is overall not beneficial as it is not cost-effective and generates misleading results and treatment. Other problems may arise when taking lymph node sections, as the extent of the

disease varies between tissues and it can be present in various stages in various cells.

1.2.8. Clinical Need

The current diagnostic techniques are unable to diagnose haematological malignancies independently. Morphological analysis must be combined with other secondary techniques in order to generate accurate and specific results. This can be expensive, as multiple techniques and medical professionals are required and it can be inefficient, as a small percentage of cases still remain undiagnosed.

There is a diverse range of leukaemia and lymphoma subtypes, which require individual and specific treatment plans, making the process complex. Thus misdiagnosing a blood cancer can cause major setbacks when providing a treatment plan to a patient. Treatments for haematological malignancies are dependent on the type and stage of the cancer. Currently, for early stage Hodgkin lymphoma, short-duration chemotherapy is combined with a low dose of involved field irradiation, which is sufficient to cure the disease (Santoro *et al*, 2006). Treatment for more advanced haematological malignancies can be very invasive, such as chemotherapy (Santoro *et al*, 2006) or bone marrow transplant, which has been adopted since the early 1980's (Burnett & Eden, 1997, Kuby, 1997 & Hoffbrand *et al*, 2010). Approximately 50 % of patients with advanced stage lymphomas relapse following high dose chemotherapy treatment (Santoro *et al*, 2006). The earlier the cancer is detected and correctly diagnosed, the earlier it can be treated and thus the type of treatment required should theoretically be less demanding on the patient. It therefore follows that the earlier the diagnosis the better the prognosis as the treatment can be provided prior to the cancer advancing to an untreatable level. If treatments such as chemotherapy can be avoided or the need for them reduced, a patient's way of life can be improved. There is therefore a need to replace the current "gold standard" techniques with methods that are less time consuming, more accurate,

and less invasive with the potential to be overall cost effective with a smaller combination of techniques being required.

1.3. Optical Diagnostic Methods

Recent research has explored the potential of optical methods as new tools in clinical diagnosis. Such methods include fluorescence, Raman and infrared spectroscopy, which utilise radiation from the electromagnetic spectrum to probe molecules and obtain information regarding biochemical composition. These methods have the potential to be rapid, non-invasive, and provide highly accurate real-time information thus eliminating the need for histological and surgical methods that are currently used in diagnosis. Diagnoses can therefore be made at a faster rate, with treatment provided earlier, reducing the emotional distress caused to patients and potentially the costs involved (Bigio & Mourant, 1997).

1.3.1. The Electromagnetic Spectrum

In 1800 William Herschel discovered infrared (IR) light when he observed a temperature change from visible to red light, and throughout the rest of the 19th Century several scientists continued to explore the nature of light to develop the electromagnetic spectrum (EMS) (Ball, 2007). The EMS includes all wavelengths of electromagnetic radiation (light) ranging from very short gamma rays up to very long radio waves (Figure 1.7). The infrared region lies in between the visible and microwave regions. It is split into near, mid and far infrared and covers the range of wavelengths from 10^{-7} to 10^{-4} m (Ball, 2007 & Smith & Dent, 2008). Both FTIR and Raman spectroscopy can utilise infrared light in order to excite molecules. The ultraviolet (UV) region of the electromagnetic spectrum is also utilised in fluorescence spectroscopy.

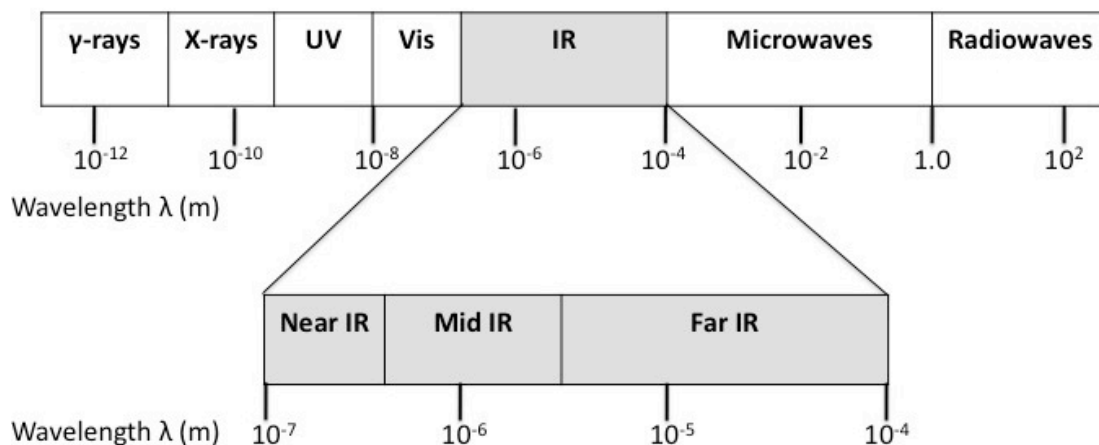


Figure 1.7. The Electromagnetic Spectrum of Radiation, the infrared region is highlighted in grey

1.3.2. The Interaction of Light with Matter

The interaction of light with matter can cause a photon (a light particle) to be emitted, absorbed or scattered. Spectroscopy is the study of spectra generated by this emission, absorption or scattering (elastic and inelastic) of electromagnetic radiation by a molecule, which is undertaken to determine the molecular composition of a sample. *Emission spectroscopy* is the molecular transition from a state of high energy to a state of low energy resulting in the emission of a photon. *Absorption spectroscopy* measures the incident radiation at varied frequencies. Upon radiation, both absorption and emission occur, and the difference (or net absorption) measured by a detector. Examples of molecular absorption of electromagnetic radiation are infrared and fluorescence spectroscopies. When electromagnetic radiation is scattered there are three different types of radiation that can be emitted: Stokes Raman, anti-Stokes Raman and Rayleigh radiation. Both Stokes and anti-Stokes Raman radiation are a form of *inelastic scattering*. In Stokes radiation, photons and molecules collide and energy is released. Lower frequency radiation is then emitted from the sample as a result of a loss in photon energy. In anti-Stokes radiation the photons gain energy from already excited molecules thus causing a higher frequency radiation to be emitted from the sample. In Rayleigh radiation there is no change

in frequency when molecules and photons collide, this is a form of *elastic scattering*. (Griffiths & de Haseth, 2007, Smith & Dent, 2008 & Atkins & de Paula, 2010).

A molecule in space can interact with photons in a variety of ways. This is due to the vibrational motions of the chemical bonds produced when molecules absorb infrared light. Linear molecules have $3n-5$ vibrational modes and non-linear molecules have $3n-6$, where n accounts for the number of atoms in a molecule (Griffiths & de Haseth, 2007). Different types of vibrational motions can occur, such as *bending*; which can entail scissoring, rocking, wagging or twisting, and *stretching*; which can be symmetric or asymmetric vibrations of the molecular bonds. In vibrational spectroscopic techniques, the light illuminating the sample alters the vibrational mode of a molecule. The change to the illuminating light is measured and plotted against the wavelength of the light used in order to produce a spectrum. This identifies the molecular entities found within a sample and is referred to as a unique biochemical fingerprint. Figure 1.8 illustrates the energy changes that arise in the different vibrational states when light interacts with matter.

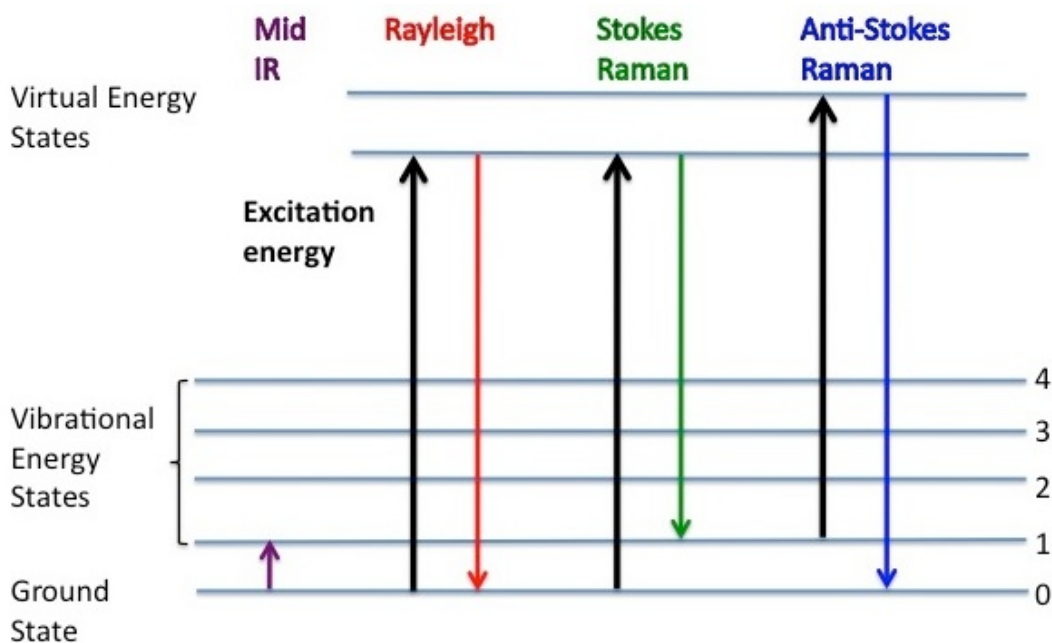


Figure 1.8. Energy Level Diagram of Mid-IR, Rayleigh, Stokes and Anti-Stokes Raman

1.3.3. Infrared Absorption Spectroscopy

Infrared absorption spectroscopy utilises mid-infrared radiation from a polychromatic light source in order to probe the molecular vibrations within molecules. These vibrations are not random but specific to the biochemical constituents of the molecule and are subject to influences from interactions with other bonds and biochemical groups as well as the surrounding environment. The information obtained from this technique allows the biochemical structure to be identified from an infrared spectrum. This spectrum displays the frequencies of the vibrational modes, which in turn provides a characteristic signature of the molecule and its biochemical structure. Infrared spectroscopy is a rapid, non-destructive and reproducible technique that is valuable in the clinical environment particularly in disease diagnosis. It has been utilised as a biomedical application in the diagnosis and assessment of the prognosis of many cancers (Ramesh *et al* 2002, Ramesh *et al*, 2003, Sahu & Mordechai, 2005, Ellis & Goodacre, 2006 & Sahu *et al*, 2006).

This method, although able to characterise the complex structures of molecules independently, is often used in combination with other complementary techniques in order to gain a broader perspective on the structure of compounds. It is also inhibited by water, which limits *in vivo* study; therefore it is commonly used alongside Raman spectroscopy, as water does not influence the spectra as strongly. These techniques also have sensitivities to different functional groups allowing a broader knowledge of the biochemical constituents of the compounds (Ellis & Goodacre, 2006, Babrah *et al*, 2007, Griffiths & De Haseth, 2007, Babrah *et al*, 2009 & Siebert & Hildebrandt, 2008).

1.3.4. Fluorescence Spectroscopy

Tissue contains a variety of chromophores, which can be fluorescent or non-fluorescent in nature. Fluorescent chromophores, such as collagen and elastin, emit fluorescence following absorption of light of a different wavelength. Non-fluorescing chromophores, such as haemoglobin have the ability to absorb the incident light without fluorescing, which can cause a reduction in energy in the spectra (Bigio & Mourant, 1997). Aside from the aforementioned endogenous fluorophores found within tissue that are fluorescent in nature, fluorescence can also be induced in tissues using precursor molecules. Commonly, 5-aminolevulinic acid (ALA) is used to induce the fluorophore protoporphyrin IX (PpIX). Fluorophores, such as fluorescein and Hematoporphyrin derivative (HPD), can also be administered in the form of drugs; these are generally referred to as exogenous fluorophores (Wagnières *et al*, 1998).

Often UV and visible light are used to excite the molecules in Fluorescence spectroscopy. When subjected to the incident light, fluorophores become excited and emit their characteristic fluorescence. Differences in the fluorescence emitted in cancerous tissue, compared to normal, are due to changes in the structure and or concentration of the absorbing fluorophores. Fluorescence spectroscopy has been utilised in the diagnosis of several cancers including breast, lung (Alfano *et al*, 1987) and oesophageal (Endlicher *et al*, 2001), thus highlighting its diagnostic potential in other cancers (Choi *et al*, 2007).

Advantages of this technique are that it has the ability to be used *in vivo*, unlike IR which is subject to interference from water. Also fluorophores such as HPD are sensitive to cancerous tissue, providing great visualisation of the contrast between tumour and surrounding tissue in the observation of cancer during treatment (Choi *et al*, 2007). However, HPD has the ability to interact with normal tissue thus it has been known to generate false positives and has a lower rate of sensitivity and specificity compared to other similar techniques such as IR and Raman (Alfano *et al*, 1984).

1.3.5. Elastic Scattering Spectroscopy

Elastic scattering spectroscopy (ESS) uses the UV/visible region of the electromagnetic spectrum covering the wavelengths from 300-900 nm. Light is directed towards a sample and once it interacts with tissue, back scattering of the photons occurs, resulting in no change to the wavelength (Anandasabapathy, 2008). The way in which light becomes scattered in tissue depends on both the structure of the tissue and the wavelength of light directed towards it. The incident light may be subject to multiple scattering episodes before being detected (Swinson *et al*, 2006). In such a case, information can be obtained from deeper tissue. It can also undergo single scattering events, which are a less common and is more often referred to as light scattering spectroscopy (LSS). In LSS, morphological information is generally only obtained from the surface of tissue. The type and number of scattering events in ESS is dependent on the scattering particles having a similar size to the wavelength of the photons (Bigio & Mourant, 1997).

The signal obtained from ESS is significantly more intense than that from other methods such as Raman and fluorescence spectroscopy. This is due to the increased number of collisions that occur between the photons and the scattering particles and thus allows the technique to be more rapid and potentially a more cost-effective method in the diagnosis of cancer (Jerjes *et al*, 2004 & Swinson *et al*, 2006). However, ESS differs from the other optical methods discussed here in that it provides structural and morphological

information from tissue rather than biochemical due to the fact that it is sensitive to changes in size and structure (Bigio & Mourant, 1997, Jerjes *et al*, 2004 & Swinson, 2006).

1.3.6. Inelastic Scattering Spectroscopy (Raman)

Inelastic scattering spectroscopy (ISS), as its name suggests, is the inelastic scattering of photons from a monochromatic light source. The inelastically scattered photons have a different energy to the incident photons as a result of the transition of the molecules from one energy state to another (Siebert & Hildebrandt, 2008). This change in energy is referred to as a Raman shift and is complementary to an IR absorption spectrum as the spectra have peaks of similar shapes and positions relating to the molecular groups (Ellis & Goodacre, 2006).

Raman spectroscopy is a very sensitive and accurate method, however the Raman shift is extremely weak due to the photons undergoing approximately only one in $10^6 - 10^8$ scattering events (Ellis & Goodacre, 2006 & Swinson, 2006). Due to this, longer irradiation times are required in order to enhance the Raman signal. This can cause over-heating and damage to the sample, but it can be overcome by having shorter exposure times and repeated accumulations of scans (Lin *et al*, 2007). Another factor that can affect Raman spectra is that organic tissue can be subject to interference from fluorescence when excited by light in the visible region of the electromagnetic spectrum. This has been overcome by the use of either near IR light or deep UV light. However, UV light has a tendency to induce photo-degradation thus near IR is the better source of radiation to use (Krafft & Sergo, 2006).

The advantages of using Raman spectroscopy alongside IR are that it is not subject to interference from water as water has a weak scattering property, thus it has the advantage that it can be used for fluid samples or even *in vivo*. It also has higher specificities to certain biochemical groups that IR does not and thus the benefit of using the techniques together allows for a broader range of

information to be obtained. (Krafft & Sergo, 2006, Lin *et al*, 2007 & Swinson, 2006.)

1.3.7. Conclusion

There are several advantages in using optical techniques for the diagnosis of cancers. Firstly, using histological methods where the morphological changes to tissue are detected is time consuming and the samples analysed are not always a true representative of the pathology. Optical techniques allow you to collect biochemical information rapidly and have the potential to be minimally invasive. In using these techniques rather than the current “gold standard” methods, it is anticipated that costs can be reduced owing to the speed of diagnosis and the reduced number of medical professionals required. By decreasing the need for biopsies and reducing the waiting time for a diagnosis, the emotional trauma to the patient can be reduced as treatment can be provided within a shorter period of time.

Although all of the methods mentioned have been utilised in cancer diagnosis they have different advantages and disadvantages to one another:

- Infrared spectroscopy is hindered by the interference of water thus *in vivo* experimentation is almost impossible, however this disadvantage can be overcome by using it in combination with other techniques such as Raman. Although withdrawing blood from patients is invasive, it is less so than other biopsy or surgical methods.
- In fluorescence spectroscopy the administration of a fluorescent drug allows it to be used *in vivo*; however this advantage renders the technique invasive and may result in side effects (Bigio & Mourant, 1997). The low sensitivity and specificity of this method also renders it less useful as a diagnostic tool for leukaemia and lymphoma than the other optical techniques described.

- Although elastic scattering spectroscopy is relatively inexpensive and can be utilised *in vivo*, the main disadvantage of it is that it does not probe the biochemical information of the sample, only the morphological changes that occur (Swinson *et al*, 2006).
- Raman spectroscopy is an extremely sensitive technique that is hindered by its weak signal and fluorescence interference. However, increasing the irradiation time of the laser and using near IR as a light source can overcome these inefficiencies.

Raman and IR are the methods of choice in this thesis owing to their complementary nature; they have the ability to obtain a wider range of biochemical information from samples owing to their specificities toward different functional groups. Raman can also be used to overcome water interference in IR, whilst IR can be used to overcome the weak signal in Raman. Recently more studies have utilised these techniques in combination, particularly in cancer (Sahu & Mordechai, 2005, Ellis & Goodacre, 2006, Krafft & Sergo, 2006, Krishna *et al*, 2006, Swinson *et al*, 2006, Lin *et al*, 2007, Isabelle *et al*, 2008 & Kendall *et al*, 2009).

1.4. Aims and Objectives

This is a continuation of a previous PhD undertaken by Jaspreet Babrah (2009) who analysed the different blood fractions from CLL samples prepared via cytospin centrifugation. This thesis compares the performance of the cytospin centrifugation samples with Drop Coating Deposition Raman Spectroscopy (DCDRS) in diagnosing leukaemia, as it is a much quicker method of depositing whole blood samples onto slides. This method, along with FTIR and Raman are explored further in chapter 2.

With the help of multivariate statistical analysis, the main aim of this thesis is to identify the performance and application of Raman and FTIR spectroscopy as a diagnostic tool for leukaemia and lymphoma and by comparing healthy samples

to cancer samples the biochemical changes that take place during cancer can hopefully be detected.

This thesis aims to achieve these aims with the following objectives:

1. DCDRS

- Determine the optimal parameters for the DCDRS method using blood proteins and extending it for use in FTIR.

2. Leukaemia

- Assessing the value of different blood fractions prepared by two different methods (cytospin centrifugation and DCDRS) in leukaemia diagnosis.
- Identifying the biochemical differences between normal and leukaemia samples by interpreting Raman and FTIR spectra of the different blood fractions.
- Developing classification models using multivariate statistical analysis and determining the performance of these models in diagnosing leukaemia.

3. Lymphoma

- Assessing the ability of FTIR and Raman to obtain spectra from fine needle aspirate biopsies from normal and lymphoma samples.
- Identifying the biochemical differences between normal and lymphoma samples by interpreting Raman and FTIR spectra.

2. RAMAN AND FOURIER TRANSFORM INFRARED SPECTROSCOPY

The use of vibrational spectroscopies for the analysis of biological materials has become a well-established tool for biochemical analysis (Mantsch *et al*, 2002, Ellis & Goodacre, 2006, Filik & Stone, 2008, Filik & Stone, 2009, Kendall *et al*, 2009). In recent years optical methods, such as Raman and Fourier Transform Infrared (FTIR), have successfully been used in the detection of several diseases, including a variety of cancers. These techniques look at changes to biochemical composition, such as protein, carbohydrate, nucleic acid and lipid make-up and aim to identify particular markers of the different stages of cancer (Dukor, 2002). The detection of cancer according to the biochemical changes allows them to be identified prior to morphological changes being apparent, thus having the potential to provide an earlier diagnosis (Kendall *et al*, 2009).

Numerous studies have been performed into the use of FTIR and Raman, which show them to be viable techniques for distinguishing the stages of blood cancers. These techniques may also be used to deduce the sensitivity of cancer treatment (Andrus, 2006). They are more reliable and accurate than current diagnostic methods as well as being more rapid and significantly less invasive. With these advantages in mind, it is hoped that these new technologies will change and improve the way in which blood cancers are diagnosed, thus aiding with prognosis and treatment for cancer patients. This chapter reviews Raman and Fourier Transform Infrared spectroscopy and their biomedical applications as well as the use of chemometrics for data analysis.

2.1. Fourier Transform Infrared Spectroscopy

FTIR is an absorption spectroscopy technique that utilises the mid-infrared (MIR) region of the electromagnetic spectrum and studies the interaction of infrared light with a sample. A range of frequencies of infrared light are directed towards a sample and absorption takes place when the frequency of the incident radiation is equal to that of a molecule's vibrational mode. Due to the absorption

of the incident radiation by the sample, the loss of frequency from the radiation is detected. What is generated is a unique 'biochemical fingerprint' of the functional groups present, owing to the induction of molecular vibrations. Information can be obtained from molecular vibrations across the entire infrared (IR) spectrum with most techniques focussing on the 800-2000 cm^{-1} wavenumber range owing to the presence of strong absorption bands of functional groups in this region. Biological material can therefore be identified by its biochemical fingerprint making it possible to differentiate between healthy and cancerous tissue according to its composition (Babrah *et al*, 2007).

In practise, MIR light is passed from a light source to an interferometer. The interferometer used in an FTIR spectrometer was designed by Michelson in the late 19th century (Figure 2.1). It acts as a beam splitter as it divides the infrared light into two separate beams of differing path length. One beam is reflected towards a fixed mirror and the other is transmitted to a moving mirror before being recombined back at the beam splitter. The purpose of this is to produce interference due to the difference in path length of the two beams generated by the differing mirrors. Once the beams have recombined it is passed through the sample and Fourier transform information about the sample is detected before being converted to a spectrum by a computer (Griffiths & de Haseth, 2007). The spectrum is a plot of the wavenumber against the absorption intensity, or percentage (%) transmittance (T). Transmittance is a ratio of the incident light (I_0) to the radiated light that comes out of the sample (I), (equation 1) which is converted to absorbance (A) using equation 2:

Equation 1:
$$T = \frac{(I_0)}{(I)}$$

Equation 2:
$$A = 2 - \log_{10} \%T$$

In order to eliminate any background interference that may be produced by the instrument or the substrate upon which the sample is set, a background

spectrum is taken from a clear area of the substrate, i.e. where there is no sample. This is then removed from the sample spectra.

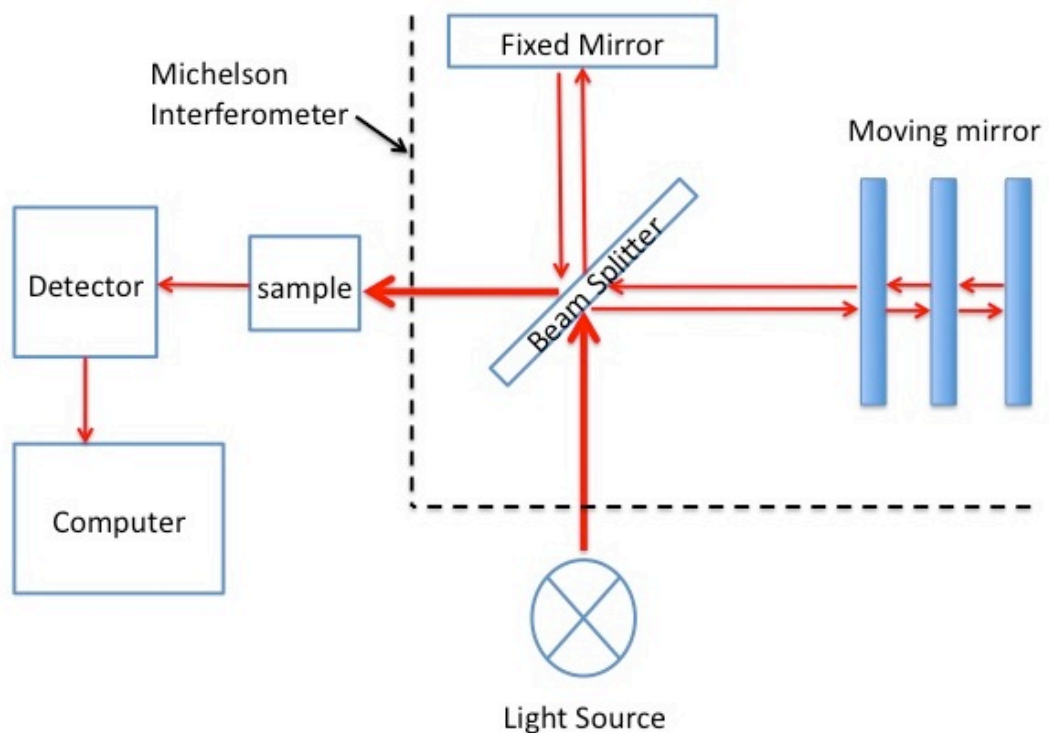


Figure 2.1. Structural Layout of a Common FTIR Spectrometer; the Michelson Interferometer is indicated in the dotted area

Infrared spectroscopy has improved greatly in the last forty years. Its primary purpose was to determine molecular structures with the aid of a monochromator, however since the introduction of the Michelson Interferometer and thus the FTIR Spectrometer in the late 1960's, the technique of structural elucidation has become significantly quicker, cheaper, and easier to use (Griffiths & de Haseth, 2007). A monochromatic infrared spectrometer simply uses a monochromator, either a prism or a grating, to split all wavelengths of the infrared radiation thus allowing only one wavelength at a time to pass through to the detector. This therefore allows measurement of only one wavelength at a time. An interferometer however allows the user to scan all

wavelengths simultaneously thus reducing the time needed to scan an entire set of wavelengths from minutes to seconds. This is known as the Fellgett Advantage. Another advantage of using an interferometer is that more energy is allowed to pass through, as there is no monochromator to narrow down its path. More energy means higher sensitivity and better quality spectra. This is known as the Jacquinot Advantage (Griffiths & de Haseth, 2007). A disadvantage of FTIR is the interference of water, particularly when analysing tissue samples, as 80 % of the body is composed of water. This can however be overcome in most cases by simply allowing a sample to dry on a slide before analysing it. Problems arise however when *in vivo* analysis is required as water cannot be removed from inside the body.

2.1.1. Biomedical and Haematological Applications of FTIR

The advantages of FTIR in the field of cancer detection are plentiful. Not only is it a reagent free, rapid, and non-invasive method, additionally it can be automated, objective and highly sensitive when combined with computational methods. It has also been proven to be successful in diagnosing other forms of cancer and therefore has the potential to be developed in blood cancer diagnostics as the new “gold standard” method (Sahu & Mordechai, 2005).

FTIR has been utilised and shown to have potential in the study of disease in tissue, single cells as well as biofluids. FTIR imaging of tissues allows rapid spectral collection from large areas through simultaneous imaging of a sample. By taking a spectral measurement at each pixel, a map of the tissue sample can be obtained. This method has been applied to the successful detection and characterisation of changes in cancerous tissues including cervical (Chiriboga *et al*, 1998b), liver (Diem *et al*, 2000), lung (Krafft *et al*, 2008), breast (Fabian *et al*, 2003) and lymph (Romeo & Diem, 2005). Single cell FTIR spectroscopy has also been utilised in the identification of the spectral differences in deoxyribonucleic acid (DNA), ribonucleic acid (RNA), phospholipids and proteins in proliferating cells (Boydston-White *et al*, 1999, Diem *et al*, 2002) including oral mucosal cells (Romeo *et al*, 2006), fibroblast cells (Mourant *et al*, 2003b) and leukaemia cells

(Schultz *et al*, 1996, Boydston-White *et al*, 1999, Boydston White *et al*, 2006). Boydston-White *et al* (1999 & 2006) studied myeloid leukaemia HeLa cell lines and identified that at different stages of the cell cycle spectral changes could be seen in the DNA, amide I and amide II protein structures due to the biochemical differences that exist when cells undergo maturation, differentiation and proliferation. Other studies have shown that dried biofluids such as whole blood (Low-Ying *et al*, 2002 & Shen *et al*, 2003), serum (Shaw *et al*, 1998, Petibois *et al*, 1999 & Liu *et al*, 2002) and urine (Shaw *et al*, 2000) can also be measured by FTIR, which shows great potential for FTIR to be used in blood cancer diagnostics, particularly for blood and serum analysis.

Babrah *et al* (2007 & 2009) recently carried out highly reproducible studies into the use of infrared spectroscopy to diagnose leukaemias and lymphomas. In a small preliminary study using lymphoma, lymphoid and myeloid cell lines they successfully identified FTIR as a viable technique for diagnosing haematological malignancies. FTIR with the aid of multivariate statistical analysis was shown to be able to identify different cancers according to very small spectral differences. A principal component Analysis (PCA) followed by linear discriminant analysis (LDA) classification model was used, the efficiency of which was validated using leave one out cross validation (LOOCV). A total of 83.6 % of the data was correctly classified with the methods achieving between 80-100 % sensitivities and specificities. The cell lines used were of pure cancer and thus more specific and more observable biochemical compositions, which would have been easier to distinguish than actual cells that may contain other structures and less obvious biochemical differences. One of the advantages of using cell lines was the ability to dry the samples prior to being subjected to the infrared light, which would be more difficult with *in vivo* studies. The use of multivariate statistical analysis was shown to be successful in differentiating the cancer types as the different cell lines clustered into their separate populations, highlighting the potential of using similar analyses in future investigations. However, in order for this study to show its full potential it would be advantageous for the techniques to be carried out on actual human cells and the results compared, thus indicating its viability in differentiating not only cancer and non-cancer but also the stages

of particular blood cancers. It has even been suggested by the authors that further work needs to be carried out at a cellular level in order to see if FTIR can still differentiate the cancers when using more complex compositions.

Several groups have identified Biochemical variations in the FTIR spectra of CLL samples. In a preliminary study by Benedetti *et al* (1985) differences in the DNA peak at 966 cm^{-1} were observed, which was assumed to be specific for lymphoid leukaemias. Using peripheral mononuclear cells Shultz *et al* (1996) identified this DNA peak to be a biomarker for CLL too. They also discovered that the amide I peak at 1656 cm^{-1} was more intense in the normal data compared to CLL and that this peak shifted to 1652 cm^{-1} in the CLL spectra. Erukhimovitch *et al*, 2006 discovered that peaks at 1056 cm^{-1} , 1270 cm^{-1} and 1592 cm^{-1} relating to carbohydrates, amide III and NH_2 vibrations of amino acids respectively, were significantly decreased in leukaemia samples.

FTIR microspectroscopy has been successfully utilised by Ramesh *et al* (2002 & 2003) in identifying the presence of cancer following chemotherapy in child acute lymphoblastic leukaemia (ALL) patients. The aims of these studies were to show the success of the treatment provided and to inspire the use of FTIR in other aspects of leukaemia research. Their results show that specific biochemical changes occur upon chemotherapy treatment, which can be identified in spectra owing to a reduction in the number of immature blood cells present. High levels of these immature cells are indicative of acute leukaemias. Although they were able to differentiate those who still had successful chemotherapy and those who did not, they were unable to identify biological markers to aid in the prognostic prediction of chemotherapy. Another similar study by Sahu *et al* (2006) looked at FTIR in both acute myeloid leukaemia (AML) and ALL patients. They looked at the biochemistry of white blood cells rather than blasts during chemotherapy treatment to identify if they returned to normal following successful chemotherapy treatment. It was shown that the biochemical profile of the white blood cells following chemotherapy did return to the normal state, as seen in control samples, suggesting the immune system had recovered and the patients

were in remission. They therefore successfully discovered potential biomarkers to use in the identification of a return to normalcy following chemotherapy.

Although FTIR has not been utilised in these studies as a diagnostic technique, the fact that it can be used to differentiate between patients who have had successful chemotherapy and not indicates that FTIR may potentially be used to differentiate these acute blood cancers. The majority of the work carried out into the use of FTIR in blood cancer diagnosis has looked at chronic lymphoblastic leukaemia (CLL) owing to the fact that it is a more common type of leukaemia and its slow progressive nature allows it to be more easily studied. These studies have proven successful in identifying useful biomarkers for leukaemia detection (Schultz *et al*, 1996 & 1997 & Erukhimovitch *et al*, 2006). Other FTIR- leukaemia studies have been shown to identify whether a cancer is high grade or low grade, however it has not been used in the classification of these cancers (Andrus & Strickland, 1998). It would be advantageous if this technique could be utilised for the classification of all types of blood cancers. However acute forms of the disease are more difficult to analyse as once they are diagnosed they need to be urgently treated owing to their aggressive nature. It is also difficult for researchers to obtain consent from acute patients to use them in such studies as they would not have very long to decide whether they would like to take part and many of the patients would be children.

2.2. Raman Spectroscopy

Raman spectroscopy, unlike FTIR, utilises the near infrared (NIR) region of the electromagnetic spectrum. Monochromatic laser light is directed towards the sample from a light source. The light first gets reflected off several mirrors towards an edge filter, which directs the laser light towards the sample placed on an attached microscope. Once the laser light hits the sample, photons in the incident light collide with molecules in the sample and become inelastically scattered. This leads to a change in energy of the photon and thus a change in its wavelength, which is otherwise known as a Raman shift (Kendall *et al*, 2009). Both Raman shifted light and laser light are reflected back towards the edge

filter where the laser light is reflected away but the Raman shifted light is allowed to pass. The Raman light is then passed through a slit to remove off axis light and is focussed onto the diffraction grating by a lens. The diffraction grating splits the light into its component parts where it is captured by a charge-coupled device (CCD) detector, which measures photons across the whole spectral range at once. (Smith & Dent, 2008.) A common Raman spectrometer layout can be seen in Figure 2.2.

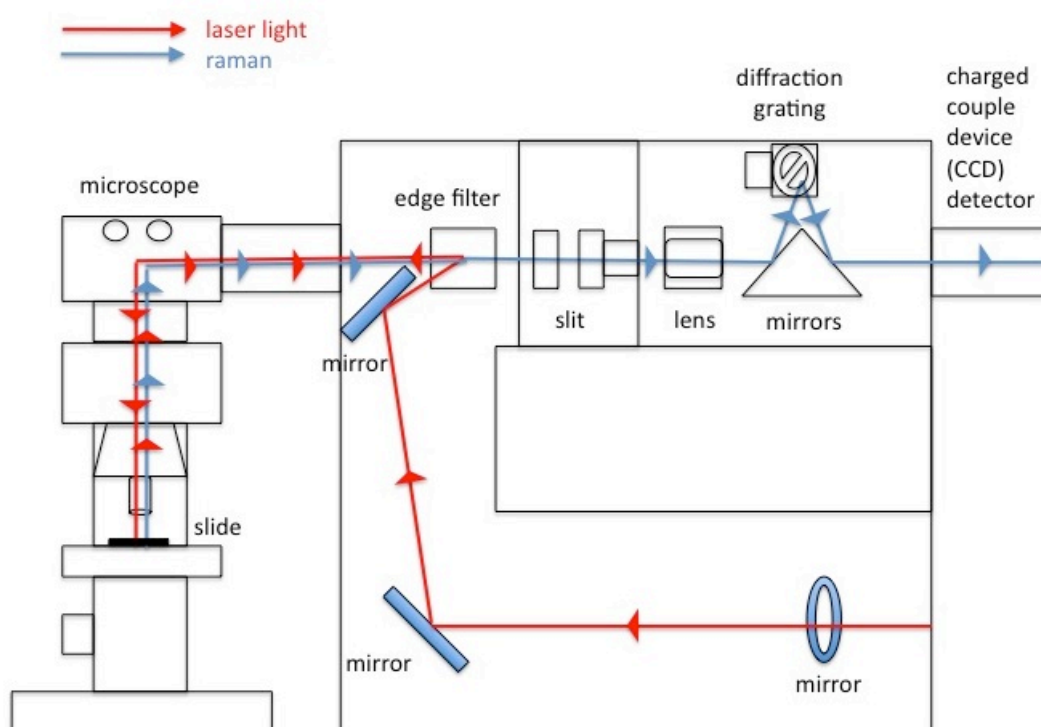


Figure 2.2. Structural Layout of a Common Raman Spectrometer

Raman spectroscopy can be used as an alternative to FTIR as well as being a complementary technique to it. The main difference between them is that in Raman a monochromatic light source is used to achieve scattering whereas polychromatic light is used in FTIR to achieve absorption. The incident light, which equals the energy required for a molecule to change from the ground state to the virtual state, causes the molecules to vibrate. As the energy level starts to fall photons in the incident light become scattered providing different

biochemical information than is generated by the absorption vibrations in FTIR. When the molecules interact with light they become polarised owing to the interaction of light with the electron clouds that encase the nuclei. This polarisation oscillates with the frequency of the illuminating photon. The molecules are said to be in an unstable virtual state and quickly return to their original state. The change of vibrational energy of the photons from the incident light is detected, unlike in FTIR where the amount of light transmitted from the sample is detected. (Ellis & Goodacre, 2006, Griffiths & de Haseth, 2007 & Smith & Dent, 2008.)

Raman can strongly identify signals that are very weak in FTIR and vice versa. Molecules with non-polar bonds such as C-C, C-H and C=C experience large changes in polarisability upon excitation which makes them Raman active (Figure 2.3A). However their dipole moment is unchanged and thus no IR absorption occurs. Polar bonds such as C-O, N-O and O-H are the opposite, having small changes in their polarisability and large changes to the dipole moment during vibration and thus are FTIR active (Figure 2.3B). These different sensitivities make the two techniques complementary (Table 2.1). (Lin *et al*, 2007, Smith & Dent, 2008 & Kendall *et al*, 2009)

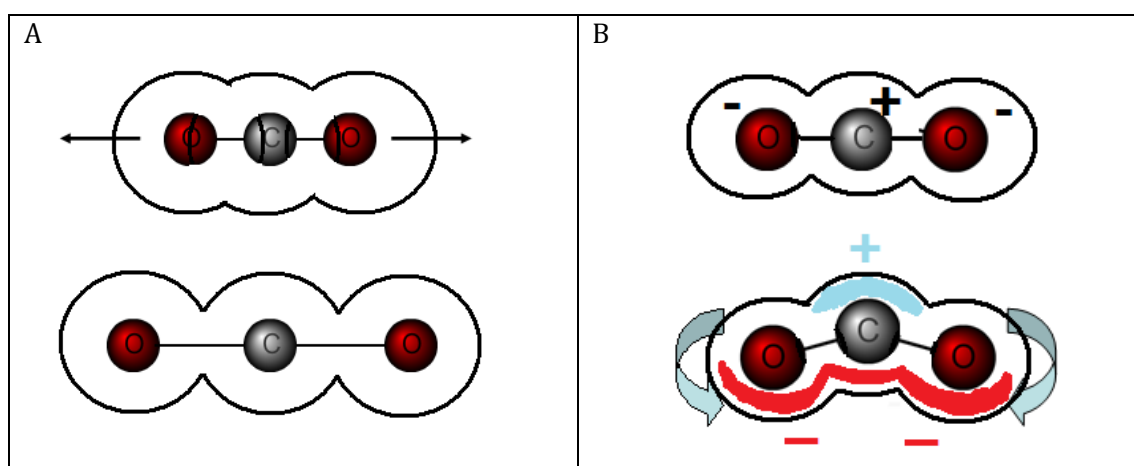


Figure 2.3. Raman and IR Active Molecules; A) Raman active molecules require a net change in the polarisability upon excitation, resulting in an increase in volume of the molecule, B) IR active molecules require a net change in the dipole moment (Source: Horsnell, 2012).

Table 2.1. Comparison of FTIR and Raman in Cancer Diagnostics

Property	FTIR	Raman
<i>Sample preparation</i>	Required	None required
<i>Wavenumber range (cm⁻¹)</i>	800-4000	400-4000
<i>Spectroscopic Detection</i>	MIR absorption with polychromatic light	NIR inelastic light scattering with monochromatic light excitation at 830nm
<i>Speed of Data Collection</i>	Fast	Slow
<i>Activated by changes in</i>	Dipole moment	Polarisability
<i>Molecular Bond Sensitivities</i>	Strong polar bonds (OH & CO) and amide bonds	Non-polar bonds (C-C double & triple bonds) and aromatic rings
<i>Signal to Noise Ratio</i>	High	Low
<i>Acceptable Substrates</i>	Mirror, (CaF ₂), barium fluoride (BaF), Low-e, zinc selenide (ZnSe)	CaF ₂ , BaF, quartz
<i>Glass Interference</i>	Strong	None
<i>Fluorescence</i>	None	Very strong
<i>Sensitivity</i>	Good	Fair
<i>Effect of Water</i>	Strong absorption and CO ₂ contribution	Weak scattering
<i>Heat Destruction</i>	None	Strong
<i>Sample Thickness</i>	Spectral saturation can occur at sample thickness of >15µm	Point spectra can be obtained at >15µm, maps at 10-20µm sections

Table adapted from Lin *et al*, 2007 & Kendall *et al*, 2009

An advantage of Raman is that it requires very little sample preparation, as water does not interfere with the spectra making it quick and simple to use. A major disadvantage of Raman spectroscopy is that only one in $10^6 - 10^8$ photons are scattered, which makes it very weak. Increased laser power is therefore required in order to increase the sensitivity of the technique and reduce irradiation time. Increasing the laser power also increases the heat that is focused on the sample, causing it to burn, irreversibly damaging it and preventing spectral information being obtained. By reducing the irradiation time and increasing the number of spectral acquisitions, the heat destruction can be minimised whilst sensitivity increased (Lin *et al*, 2007). In recent years the weak property of Raman has also been overcome with the introduction of monochromators, which increase sensitivity. There are a number of other techniques available which enhance the Raman signal, however they are currently limited to research applications and are not intended to be used in this thesis. Examples include coherent anti-Stokes Raman scattering (CARS), surface enhanced Raman scattering (SERS) and Resonant Raman Scattering (RRS) microscopy which are much more sensitive than traditional Raman spectroscopy (Burkacky & Zambusch, 2008 & Naumann, 2008). CARS spectroscopy uses a higher frequency wavelength, minimising the interference from fluorescence (Chan, 2009 & Hutchings, 2009). SERS enhances the Raman signal by utilising gold or silver particles upon which samples are adsorbed. This allows it to detect single molecules (Hutchings, 2009 & Horsnell, 2012). RRS allows photons to be scattered up to 10^6 times more than in traditional Raman spectroscopy. This is because the incident light corresponds to the energy required for an electron to be excited to a virtual state (Horsnell, 2012). The increased sensitivity of these techniques allows for both faster and more efficient real time analysis. For a more detailed description of these techniques please refer to Lasch & Kneipp (2008) or Smith & Dent (2008).

The choice of light source can also be a problem in Raman spectroscopy. For example, the depth at which ultraviolet (UV) light can penetrate a sample is much lower than NIR owing to its shorter wavelength. Thus the focus of UV light on the surface of a sample is more intense and may cause heat destruction

(Dukor, 2002 & Lin *et al*, 2007). Another disadvantage of Raman spectroscopy is fluorescence interference, particularly in tissue. This is due to the fluorescing nature of certain cellular components when subject to visible light. Since fluorescence is reduced more quickly at longer excitation wavelengths NIR light can be used as an alternative in order to minimise this (Dukor, 2002, Griffiths & de Haseth, 2007 & Smith & Dent, 2008). Selection of the excitation laser wavelength (λ_0) therefore depends on several factors. The Raman differential scattering cross-section (s' (cm²/sr)) varies inversely with the fourth power of the excitation wavelength by the proportionality formula as shown in equation 3:

Equation 3: $s' \propto (\lambda_0 - \lambda_{vib})^{-4}$

Where: λ_0 is the wavenumber of the incident radiation ($1/\lambda_0$)
 λ_{vib} is the wavenumber of the vibrational mode (λ_{vib}/c)
 c is the speed of light

Selecting a shorter laser wavelength can increase the Raman scattering cross-section. However, as laser energies reach into the shorter visible and UV wavelength range, fluorescence can become pronounced, which may interfere with weaker Raman signals. In addition, the Raman shift is constant with regard to vibrational energy, hence constant in wavenumber, but not wavelength. A greater Raman shift makes separating the elastically scattered and Raman scattered light much easier, possibly resulting in improved signal-to-noise ratios and the ability to measure smaller Raman shifts. Thus, the Raman excitation wavelength should be selected in consideration of the Raman scattering cross-section, the magnitude of the Raman shift in wavelength, and the potential for fluorescence.

2.2.1. Biomedical and Haematological Applications of Raman

Raman spectroscopy has been utilised in the early detection of several cancers including lymph node, gastrointestinal, lung, cervical, brain and breast cancer (Krafft & Sergo, 2006, Lyng *et al*, 2007, Taylor *et al*, 2007, Kendall *et al*, 2009, Orr

et al, 2010). It would be advantageous if it could also be used in the diagnosis of blood cancers alongside FTIR owing to its complimentary nature and the fact that FTIR has already been successfully utilised for this purpose. Little work has been carried out into the use of Raman spectroscopy of cells compared to FTIR and thus Krishna *et al* (2005) carried out some preliminary work with mixed cancer cell lines showing its potential to identify the different spectral features that exist in the different cancers. One small study recently carried out by Kendall *et al* (2009) on myeloid leukaemia cell lines indicated the feasibility of Raman in identifying the biochemical compositions of the cancer. Although they have suggested that Raman could be used alongside FTIR, little information is reported into the nature of the experiment and it would appear that a vast amount of work is still required. Looking at other research undertaken into Raman spectroscopy with lymph nodes would be quite informative as to the ability of Raman in leukaemia and lymphoma diagnosis. Orr *et al* (2010) have recently studied lymph nodes and successfully compared the results obtained through Raman spectroscopy of Hodgkin and non-Hodgkin lymphoma. Using their model to predict nodes with and without cancer they obtained over 80% accurate spectral prediction, again highlighting the potential of Raman spectroscopy in replacing the current “gold standard” morphological techniques. As with FTIR, Raman Spectroscopy has been exploited for tissue, cell and biofluid analysis. Pully *et al* (2010) obtained 32 x 32 pixel images of single living peripheral blood lymphocytes with collection times of two minutes using time-lapse Raman imaging (TLRI). Other Raman imaging studies of cancerous tissues include lung (Krafft *et al*, 2008), cervical (Mantsch *et al*, 2002), skin (Mantsch *et al*, 2002) and brain (Kendall *et al*, 2009). Raman has also been used for the analysis of dried teardrops (Filik & Stone, 2008, Filik & Stone, 2009) urine (Reyes-Goddard, *et al*, 2005), plasma (Poon *et al*, 2012) and whole blood (Enejder *et al*, 2002 & Reyes-Goddard, *et al*, 2005) highlighting the potential for using whole blood in leukaemia diagnosis. Many groups are now using both Raman and FTIR alongside each other to probe further biochemical information from samples for diagnosis (Mantsch *et al*, 2002, Ellis & Goodacre, 2006, Lyng *et al*, 2007, Krafft *et al*, 2008, Filik & Stone, 2008, Filik & Stone, 2009 & Kendall *et al*, 2009).

2.2.2. Drop Coating Deposition Spectroscopy

Both Raman and Fourier Transform infrared spectroscopies are highly powerful techniques, but they have their limitations and as such new methods are constantly being developed to enhance their capabilities. Drop coating deposition spectroscopy (DCDS) has been developed to enable microanalysis of compounds. The fundamental concept of the technique has been around since Johann Andreas von Segner first described surface tension in 1751 (Young, 1805). The formation of a spherical shape from a droplet of fluid arises due to the attractive forces of cohesion; forces between the molecules in the liquid, and adhesion; forces between the molecules in the liquid and the substrate on which the droplet is formed (Kočevár, 2005). Over the past 200 years, scientists have continued to research and understand how solutes adhere to slides and how it can be utilised for biological analysis (Rowan *et al*, 1995, Parisse & Allain, 1996, Deegan *et al*, 1997, Deegan, 2000, Deegan *et al*, 2000, Kopecky & Baumruk, 2006, Filik & Stone, 2007, Esmonde-White *et al*, 2009, Sefiane, 2010).

Deegan *et al* (1997) accurately described the process of ring formation from a sessile droplet as similar to ‘when a spilled drop of coffee dries on a solid surface’, leaving behind ‘a dense, ring-like deposit along the perimeter’. A more scientific explanation is that when a drop of liquid is deposited onto a solid surface, such as a slide, it evaporates from the outside edge causing an outward flow of liquid from the interior as it tries to replenish the evaporating fluid. The outward flow carries solutes with it, depositing them at the outer edge, a process commonly referred to as pinning of the contact line. What remains is a dried ring that is highly concentrated with evenly dispersed solute (Figure 2.4). Without the deposition of the solutes, a contact line would not be formed and the droplet would shrink (Deegan, 2000, Deegan *et al*, 2000, Chen & Evans, 2010).

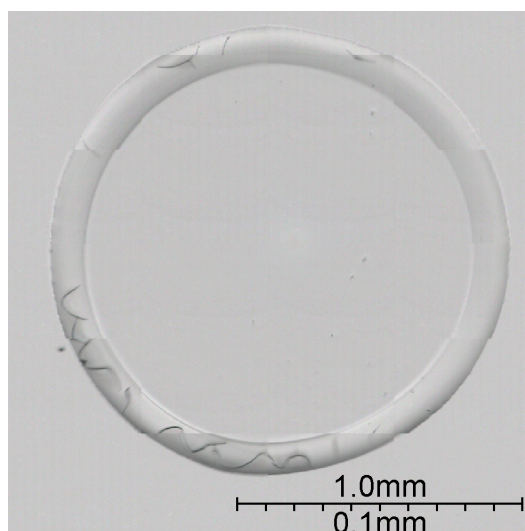


Figure 2.4. Image of a Typical Droplet Deposited on a CaF₂ Slide

Using droplets to form rings of highly concentrated solutes can be utilised for biological analysis. Raman and Fourier Transform Infrared spectroscopy (FTIR) require samples of a limited thickness and concentration for information to be obtained. By using drop drying to concentrate solutes, an increased amount of information can be obtained from low volume, dilute samples. Zhang *et al* (2003) first coined the drop coating deposition Raman spectroscopy (DCDRS) term by depositing droplets onto a slide, forming a ring pattern of concentrated solution, and obtaining protein spectra using Raman spectroscopy (Ortiz *et al*, 2006). Since then, several studies have been carried out into its use for protein analysis due to its high sensitivity and reproducibility (Filik & Stone, 2007, Ortiz *et al*, 2006).

Using a series of dilutions of the most common proteins found in human blood (albumin, fibrinogen and globulin), the DCDRS technique is explored and extended for use with FTIR in order to identify the upper and lower limits of the technique. Thus from here onwards it will be referred to as DCDS.

In chapter 5 the DCDS technique is utilised for biological analysis of whole blood samples for leukaemia diagnosis, which have unknown or uncontrollable concentrations of components. A major drawback of whole blood analysis is the

low concentration of components which have weak signals that are easily masked by the strong absorption and scattering properties of the red blood cells themselves (Enejder *et al*, 2002). Using this technique allows this problem to be overcome as the components of the blood cells spill out into solution forming rings that are evenly dispersed and highly concentrated.

2.3. Data Analysis

Interpretation of individual IR and Raman spectra can prove challenging and subjective due to the complexity of the peaks and the very small differences that arise from the various vibrational modes of molecules. Thus the use of chemometrics, or multivariate statistics, to develop classification models to identify spectral variance is vital for the analysis of the vast amounts of data that will be obtained. The techniques used in this thesis are described below.

2.3.1. Spectral Pre-Processing

Both Raman and IR spectra need to be pre-processed prior to chemometric analysis. This can involve several methods, which aim to reduce interference from background noise so that the relevant peaks are not suppressed or masked by other artefacts.

Normalisation

Normalisation removes differences between spectra due to signal intensity or absorbance. This can be achieved using an internal standard such as a peak within the spectra, however the shape and position of the peak must not be overlapped by the signals of other compounds. Alternatively it can be normalised according to the area under the curve, although this can be affected by negative values. Here, vector normalisation is used where the sum of the squared values is equal to 1.

Mean Centring

Mean centring the data involves subtracting the mean of the entire dataset from each spectrum so that the intensity of the spectra is to the same scale. This allows the differences from the mean value to be explored.

2.3.2. Peak Ratio Analysis

Peak ratio analysis was carried out in order to remove any saturated or weak IR spectra according to the ratio of the amide I and amide II peaks. In FTIR spectroscopy, samples thicker than 15 μm are more highly absorbing and thus the intensity of the light emitted from the sample is much lower. According to equation 1, this causes the %T to be very high and thus the spectra obtained is highly saturated (Kendall *et al*, 2009). More incident light is absorbed and thus the intensity of the light emitted from the sample is very low causing the %T to be very high. It is therefore not a true representation of the biochemical make-up of the sample. To eliminate these spectra, ratios of the amide I and amide II peaks were identified from various sources, and any ratios that were lower than a specified value were removed from the data sets.

2.3.3. Chemometric Analysis

Difference Spectra

Difference spectra were produced to identify the main differences between the data groups. By subtracting the mean spectra of one group from the mean of another the similarities between the groups are removed, leaving behind only the differences. The remaining peaks were then identified to look for any changes between the pathologies due to protein, DNA and other biochemical constituents. In order to identify the peaks a table of Raman peak shifts and IR peak intensities was created from various references (Appendix B).

Principal Component Analysis (PCA)

PCA is used to reduce or simplify large datasets into smaller more manageable variables, which better represent the differences in the data and thus the data as a whole. It looks at the significant peak differences and separates spectra based on their variance. Principal components (PC) are generated, with PC1 representing the largest variance, followed by PC2, then PC3 etc, with noise appearing in later PCs. The most significant PCs are then plotted against each other to identify any natural separation that occurs (Ellis & Goodacre, 2006 & Brereton, 2009). This is as an unsupervised technique.

Linear Discriminant Analysis (LDA)

The PCs generated can then be fed into LDA, which maximises the separation between the groups according to a linear function and is therefore a supervised technique. LDA is used to reduce the variance within each group whilst maximising the variance between the different groups. It therefore separates the groups according to their pathology creating a model that future data can be inserted into and its pathology predicted. The number of PCs used must be minimal in order to prevent any over-fitting of the data. PCA followed by LDA has been used by several groups for the analysis of both IR and Raman data (Kendall *et al*, 2003, Lyng *et al*, 2007, Babrah *et al*, 2009 & Babrah, 2009).

Leave One Sample Out Cross Validation (LOSOCV)

Finally, LOSOCV looks at the quality of the data and the classification method. All of the data from one sample (the test set) is removed from the model and the analysis rerun using the remaining data (the training set) to identify if the model can classify the remaining samples as accurately. This is repeated with all of the data for each sample and sensitivities and specificities are calculated and a contingency table produced. Sensitivity is a measure of the number of true positives that are correctly identified and is calculated as shown in equation 4, whereas specificity is a measure of the number of true negatives that are correctly identified and is calculated as shown in equation 5. By generating sensitivities, specificities and contingency tables (which displays the distribution

of the data across the pathology groups) it highlights the ability of the classification model to separate the groups according to their pathology.

$$\textbf{Equation 4: Sensitivity} = \frac{\textit{no.true positives}}{\textit{no.true positives} + \textit{no.false negatives}} \times 100\%$$

$$\textbf{Equation 5: Specificity} = \frac{\textit{no.truenegatives}}{\textit{no.truenegatives} + \textit{no.false positives}} \times 100\%$$

Although there is a plethora of statistical techniques available, all of these techniques are accepted and sufficient for this thesis. There may be other techniques that are more optimal but their use is beyond the scope of this work.

2.4. Conclusion

The use of vibrational spectroscopies, such as Raman and FTIR, in early cancer diagnosis has advanced significantly in the last 15-20 years (Griffiths & de Haseth, 2007). Preliminary studies have been carried out into a vast array of cancers, which have shown them to have great diagnostic potential. However, in nearly all cases more research is needed to prove that they are highly accurate and cost-effective to use as a medical diagnostic tool in replacement of the methods already available. These techniques were originally used as separate tools but they have shown great potential to be used in combination with each other. Kendall *et al* (2009) have reviewed several preliminary, combinatorial studies on FTIR and Raman in cancers such as lung, lymph, brain, breast and cervical. The success of this research has not only identified the future potential but has also aided in encouraging the techniques to be used in replacement of the current morphological ones. Krishna *et al* (2006) have used Raman and FTIR together to identify drug resistance in leukaemia cell lines thus highlighting the capability of these techniques to identify unique spectral changes that can occur within similar samples and thus characterise the status of the cells.

A key advantage of using FTIR and Raman together is that they have different sensitivities to particular molecular entities (Table 2.1). For example, FTIR is highly sensitive to polar functional groups such as C=O, C-O, C-H and O-H, whereas Raman is more sensitive to C=C and carbon triple bonds. Therefore not all the vibrations within a molecule may be both IR and Raman active; thus using the techniques together can provide a better overall biochemical composition of the molecules. Other benefits of using these techniques in combination allow the disadvantages of water in FTIR to be overcome. (Kendall *et al*, 2009 & Smith & Dent, 2008.)

Both Raman and FTIR can be used to identify the morphological as well as the biochemical structure of a sample. With the attachment of a microscope to the instrumentation they can be used to image the samples in real-time thus providing even more information than the current diagnostic “gold standard” techniques do. A vast range of biological molecules can be studied owing to the combinatorial nature of these techniques. The major components of tissues (nucleic acids, proteins, lipids and carbohydrates) have unique biochemical fingerprints, which can be easily identified in both Raman and FTIR spectra (Dukor, 2002). Along with their non-subjective, complementary nature, their speed and ease of use, they are able to provide accurate and reproducible results and have great potential to be used as the future “gold standard” techniques in the diagnosis of blood cancers. However, future work into the *in vivo* capabilities of these techniques is required in order to show that they are better than the current methods employed.

Currently an insufficient amount of research has been carried out into the potential of FTIR and Raman in diagnosing haematological malignancies. Although many research groups have successfully used them for the diagnosis of other cancers in the last ten years, more research is needed into their accuracy and viability as a medical application (Sahu & Mordechai, 2005). In some cases where controlled studies have been carried out, FTIR appears to have an 80 % prediction rate. However these results seem to be less promising in uncontrolled studies thus identifying the need for more in-depth research (Dukor, 2002).

There appears to be very little literature available on Raman and blood cancers analysis using patient blood samples, thus it has been beneficial to look at lymph node studies in order to gain some perspective on its use in haematological malignancies. Isabelle *et al* (2008) have used Raman and FTIR in combination to differentiate between cancerous and non-cancerous lymph nodes. Although they have not used the techniques to differentiate between cancers and their subtypes, they have shown potential for the techniques to be used to identify the chemical compositions of lymph node cancers such as Hodgkin and non-Hodgkin lymphomas.

2.5. Summary

In this chapter the vibrational spectroscopies FTIR and Raman and their diagnostic applications have been explored. The methods that will be used in this thesis for the diagnosis of Leukaemia and Lymphoma have also been introduced. In the following chapters these methods are described in more detail and the results of the investigations are presented.

This thesis follows on from the work carried out by Jaspreet Babrah (2009) on leukaemia cell lines. Thus her thesis should be referred to for more information regarding the initial work that was undertaken into these methods.

Unfortunately, due to the limited number of samples obtained throughout this PhD the work involving the use of lymphoma samples can only be described as a feasibility study.

3. MATERIALS AND METHODS

Analysing the Raman and Infrared (IR) spectra of four different sample types obtained from two different sample preparation methods for the diagnosis of leukaemia is the main focus of this research. The first method, Drop Coating Deposition Spectroscopy (DCDS), is explored with blood proteins in chapter 4 prior to the analysis of whole blood, buffy coat and plasma samples from leukaemia patients and healthy volunteers in chapter 5. This method has been utilised as it has the ability to gain information from a small sample volume and involves minimal sample preparation (Filik & Stone, 2007). If leukaemia can be accurately diagnosed from a small sample volume combined with a rapid analytical technique, there is potential for this method to give almost instantaneous diagnostic results to patients thus avoiding up to a two-week wait for results (Three Counties Cancer Network, 2005).

Also in chapter 5 the second method, Cytospin Centrifugation, is explored. This method is currently used in haematology laboratories to separate blood fractions (Hoffbrand *et al*, 2010). The plasma, buffy coat (containing the white blood cells) and the red blood cells can be extracted and analysed spectroscopically to obtain information across all three fractions that may not be available from the fractions individually. This method has been used previously with the buffy coat fraction achieving 99.5 % correct classification, plasma achieving 80.6 % and whole blood 70 % with Infrared Spectroscopy (Babrah, 2009). Although buffy coat has been shown to give the best results, separating the blood fractions prior to spectroscopic analysis is laborious, taking up to two hours and utilises more resources and medical professionals. Analysing whole blood and obtaining both FTIR and Raman spectra alongside may be valuable in identifying differences that may not be seen in the buffy coat spectra alone. The whole blood cytopsin sample can also be compared to the whole blood DCDS sample.

Other methods have been explored for the preparation of blood samples but are appropriate or advantageous over the methods described. For example blood

smears could be used to analyse whole blood. However, the process of obtaining a blood smear is highly specialised and takes time and practise often resulting in high wastage of slides. This makes the method expensive owing to the nature of the Calcium Fluoride (CaF_2) slides required for these spectroscopic methods and their individual cost. It therefore holds no benefit over cytopsin centrifugation of whole blood, which is a much quicker and more practical method of obtaining whole cells.

Fine needle aspirate biopsies (FNAB) were obtained for spectroscopic analysis of lymphoma. The results from the lymphoma feasibility study are displayed in chapter 6.

3.1. Optimisation of Spectral Measurements: Identifying Parameters

Preliminary measurements were carried out to determine the optimal parameters for Fourier Transform Infrared (FTIR) and Raman in order to avoid background interference or loss of data in the spectra. Background noise is an important factor to consider when measuring IR and Raman spectra of biological samples. The overall effect of the noise to the spectra is dependent on the signal to noise ratio (SNR). This is a ratio of the total signal intensity divided by the noise intensity and it is proportional to the square root of the signal intensity multiplied by the acquisition time. For example, in Raman spectroscopy if you increase the intensity of the laser light, the acquisition time can be reduced thus allowing for a more rapid measurement with better spectral resolution (Kendall, 2002). The parameters that were explored are displayed in Table 3.1.

Table 3.1. FTIR and Raman Parameter Options

System	Parameter	Range
<i>FTIR</i>	Aperture Size	10 x 10 – 100 x 100 μm
	Resolution	1 - 64 cm^{-1}
	No. scans per pixel	1 - 64
<i>Raman</i>	Objective	x5, x20, x50, x80
	Acquisition time	1 – 50s
	Accumulations	1 - 10

3.1.1. Fourier Transform Infrared Spectroscopy

The infrared spectra were produced using infrared micro-spectroscopy (microscopy): an FTIR instrument linked to a microscope. The specific system being a Perkin Elmer Spectrum One FTIR spectrometer coupled to a Spotlight 400 molecular imaging microscope. The sample, located on the centre stage of the microscope, is encased in a chamber to reduce outside influences from the surrounding environment. The system has a liquid nitrogen cooled mercury cadmium telluride, HgCdTe, (MCT) linear array detector with a 16x1 element. Comparisons of the number of scans per pixel, resolution and the aperture size were measured and the most accurate, time efficient results were noted for use in following experiments. The SNR was calculated for each of the different spectra obtained in Figures 3.1- 3.3 and is displayed in Table 3.2 below.

Table 3.2. SNR for FTIR Spectra

Resolution (cm^{-1})	SNR	No. Scans	SNR	Aperture (μm)	SNR
1	1:44	1	1:344	10 x 10	1:327
2	1:12	2	1:365	15 x 15	1:150
4	1:209	4	1:167	20 x 20	1:165
8	1:370	8	1:181	30 x 30	1:354
16	1:123	16	1:353	50 x 50	1:310
32	1:96	32	1:157	100 x 100	1:281
64	1:336	64	1:250		

Using a section of lymph node tissue on a CaF₂ slide, twenty point spectra were measured at different resolutions from 1 - 64 cm⁻¹ and the means of the spectra compared (Figure 3.1). At the higher resolutions, 1cm⁻¹ – 8 cm⁻¹, the spectrum is more defined with increased spectral detail visible compared to the lowest resolution, 64 cm⁻¹. This is evident as the peaks appear to become smoother, as the finer detail and thus spectral information is lost. Some peaks can even be seen to overlap, for example at 1650 cm⁻¹ the doublet visible at all spectral resolutions, appears as a singlet at the lowest resolution. Looking at the SNR in Table 3.2 there does appear to be some variation in noise, however this appears to be completely random and not related to the decrease in resolution. Therefore in order to obtain the best quality spectra a resolution of 4 cm⁻¹ was selected for future analyses.

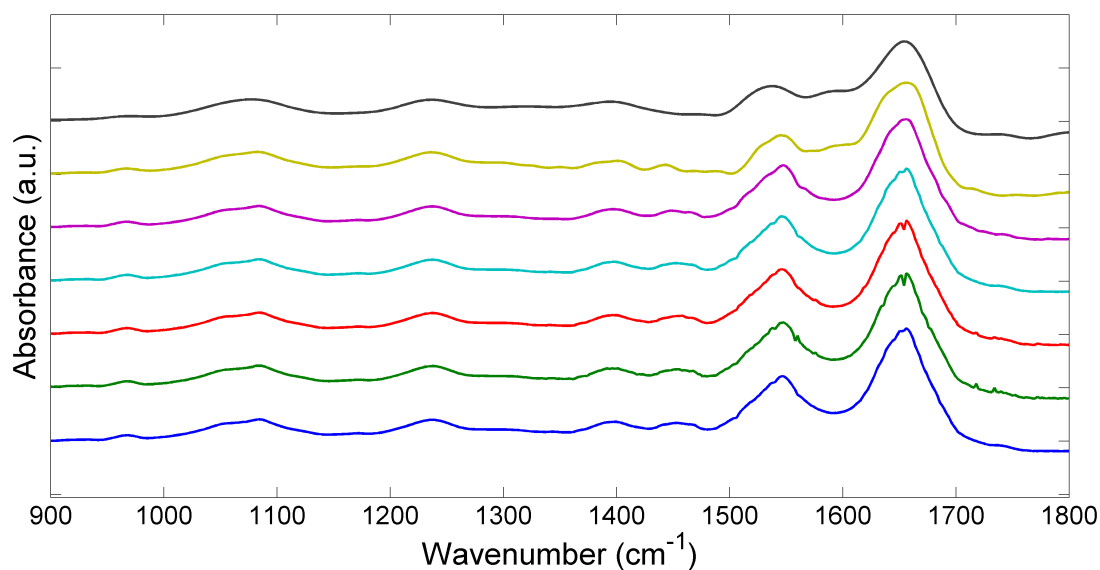


Figure 3.1. Mean FTIR Spectra of Lymph Node Tissue Sections To Identify the Effect of Increasing Spectral Resolution on Quality of Spectra; measured at 1 (blue), 2 (green), 4 (red), 8 (turquoise), 16 (pink), 32 (yellow) and 64 cm⁻¹ (black) resolution

Increasing the number of scans per pixel does not appear to vastly improve the quality of the spectra. However the more repeats that are done, the more likely it is that any random noise in the signal is smoothed out and thus an appropriate number of scans needs to be determined in order for the analysis to be time efficient. Again, twenty point spectra were obtained from a lymph node tissue section to show the change in the quality of the spectra as the number of scans run per pixel increases from 1 to 64 (Figure 3.2). As the number of scans increases the quality should theoretically improve as an average spectra is obtained. However, the difference seen between 1 scan and 64 scans does not appear to be very significant, and this is reflected in the SNR values in Table 3.2. Therefore it seems sensible to run 16 scans per pixel so as to get an appropriate average to generate enough detail, without causing the run-time to be unnecessarily high.

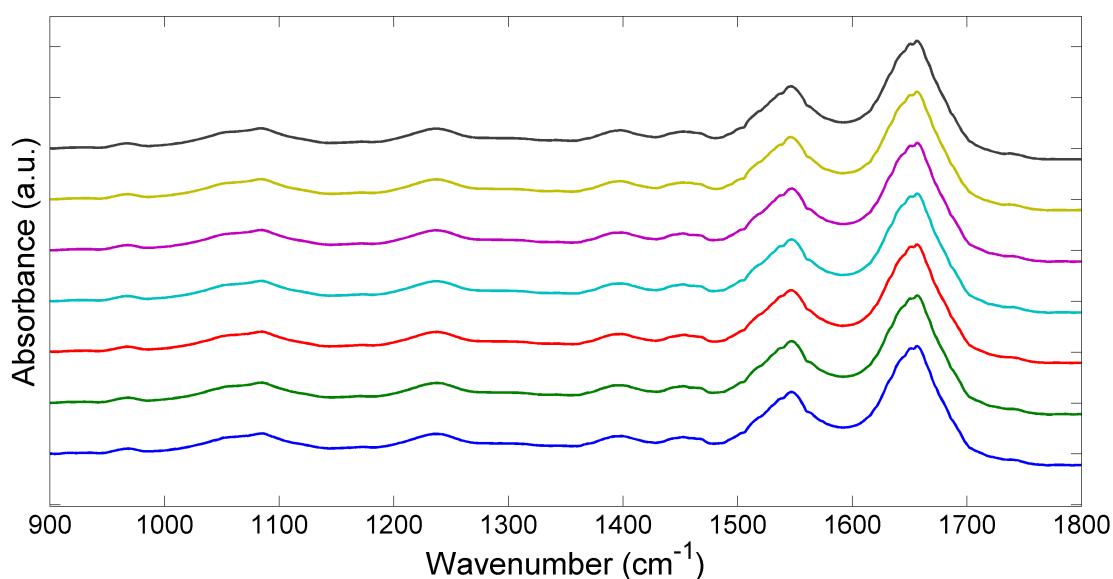


Figure 3.2. Mean FTIR Spectra of Lymph Node Tissue Sections To Identify the Effect of Increasing the Number of Scans Per Pixel on Quality of Spectra; measured at 1 (blue), 2 (green), 4 (red), 8 (turquoise), 16 (pink), 32 (yellow) and 64 (black) scans per pixel

Pixel size, or aperture, is also a factor to consider as the smaller the pixel, the more well defined the area that the spectrum is measured from as the area being focused on is much smaller. Figure 3.3 show the change in spectra as the aperture is increased from a 10 x 10 to a 100 x 100 μm square. As the aperture is reduced the spatial resolution of the spectra is changed and therefore samples a different amount of the material of interest. The spectral detail and SNR (Table 3.2) does not appear to change with aperture size and thus the 30 x 30 μm aperture was selected in order to obtain time efficient spectra.

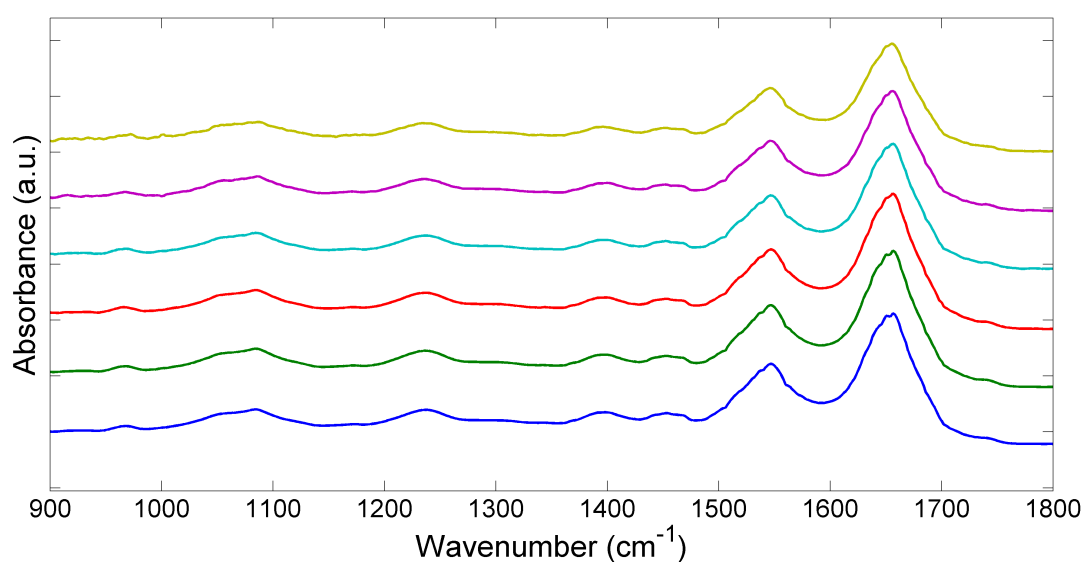


Figure 3.3. Mean FTIR Spectra of Lymph Node Tissue Sections To Identify the Effect of Increasing Aperture Size on Quality of Spectra; 10 x 10 (blue), 15 x 15 (green), 20 x 20 (red), 30 x 30 (turquoise), 50 x 50 (pink) and 100 x 100 μm (yellow) aperture size.

3.1.2. Raman Spectroscopy

Using the same lymph node tissue sample and taking the mean of twenty point spectra, similar experiments were undertaken to identify optimised settings for the Raman system. The Raman instrumentation is a customised Renishaw Streamline System 1000 micro-spectrometer that has been optimised for tissue work in near-IR (NIR). The Raman spectrometer contains an 830 nm wavelength laser, and is attached to a Leica confocal, polarised light microscope. This setup is a dispersion system, which uses a diffraction grating rather than an interferometer to split the light into its component parts. The diffraction grating has 300 lines per mm to allow a full-spectral scan without the need to move the grating. The edge-filters are metal oxide in order to optimally reject the elastically scattered light and the detector is a thermoelectrically cooled (to -70 °C) deep depletion charged couple device (CCD) detector, which reduces any thermal background noise. High quality Leica objective lenses that do not produce significant fluorescence are also selected. Comparisons of the objective, acquisition time for each spectra and the number of accumulations of each measurement were made. The SNR was also calculated for each of the different Raman spectra obtained in Figures 3.4- 3.6 and is displayed in Table 3.3 below.

Table 3.3. SNR for Raman Spectra

Objective	SNR	Acquisition Time (sec)	SNR	No. Accumulations	SNR
X 5	1:6	1	1:9	1	1:12
X 20	1:4	2	1:8	2	1:10
X 50	1:5	5	1:8	5	1:7
X 80	1:6	10	1:5	10	1:6
		20	1:2		
		30	1:2		
		50	1:2		

Raman parameters need to be adjusted to suit the sample type in order to generate good quality spectra in an efficient manner. The resolution is dependent on the objective used. The higher the objective magnification, the higher the numerical aperture, and the better the resolution. As in FTIR, high

resolution allows more localised detail to be gained from a sample. At the lower objective lens, x5, the resolution is the lowest and some of the spectral detail appears to be lost (Figure 3.4). This is due to the collection efficiency of the lens, which relates to the numerical aperture of the objective i.e. the ability to distinguish between two points. The higher the magnification objectives have a higher numerical aperture and therefore collect more light. The spectrum for the x20 objective is more defined with the peaks representing the biochemical information clearly visible. Using the x50 and x80 objectives the peaks become even more defined as more biochemical information is visible in the spectra. Although the spectra appear best at x80 magnification in this example, the objective selected will depend on the sample being analysed as the volume of interest will vary with each sample. Furthermore the Raman signal is very weak and the objective used should achieve the best possible spectra. The SNR does not appear to change with the objective used (Table 3.3).

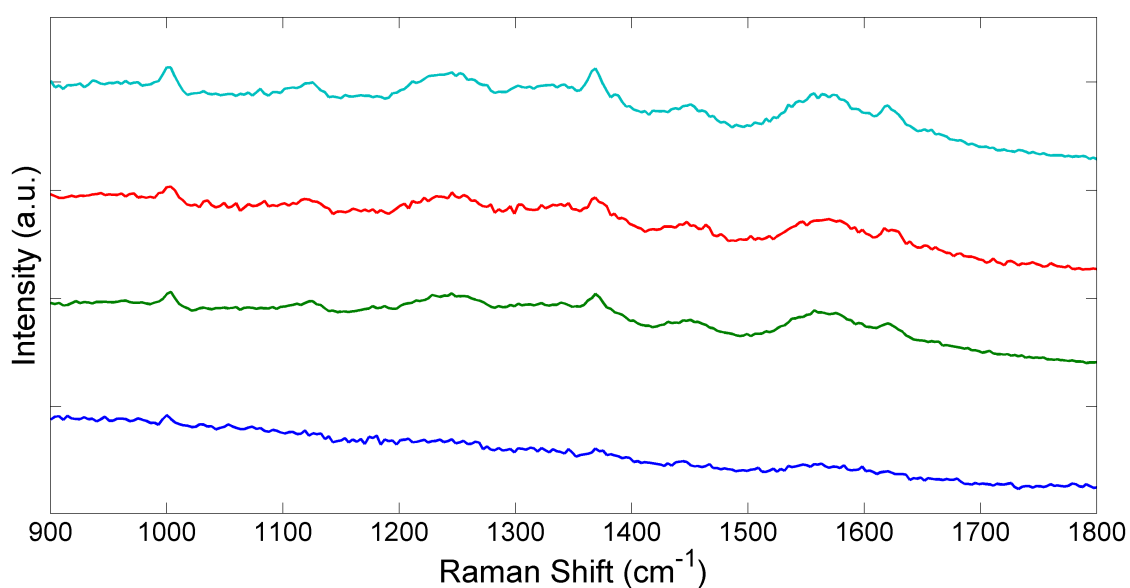


Figure 3.4. Mean Raman Spectra of Lymph Node Tissue Sections To Identify the Effect of Different Objectives on Quality of Spectra; x5 (blue), x20 (green), x50 (red) and x80 (turquoise) objective lens

The length of time the spectra are acquired for also needs to be considered, as shown in Figure 3.5. As the spectral acquisition time increases, the SNR also improves as highlighted in Table 3.3. However, the longer the acquisition time, the less time efficient the technique is and it would appear that the spectra do not vastly improve much above thirty seconds. The effect of the laser heat on the sample also needs to be taken into consideration as the longer the laser is projected onto the sample the more chance there is of the sample burning. Therefore it does not seem appropriate to acquire spectra for longer than thirty seconds.

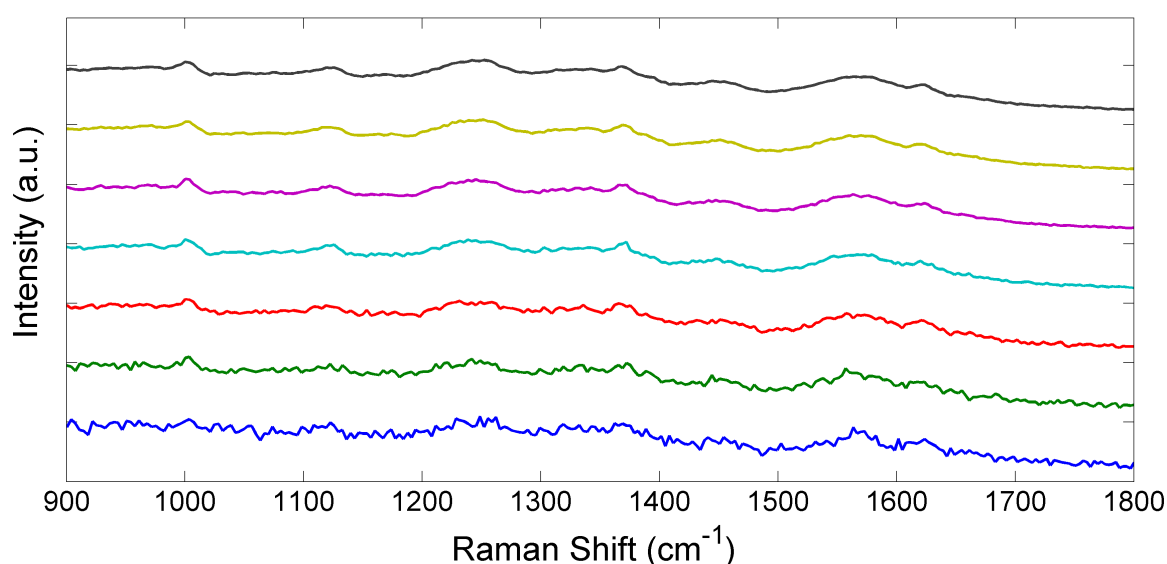


Figure 3.5. Mean Raman Spectra of Lymph Node Tissue Sections To Identify the Effect of Spectral Acquisition Time on Quality of Spectra; 1 second (blue), 2 seconds (green), 5 seconds (red), 10 seconds (turquoise), 20 seconds (pink), 30 seconds (yellow) and 50 seconds (black) acquisition time

Finally, increasing the spectral accumulations at each point on a sample increases the quality of the spectrum, as the effect of multiple accumulations is comparable to longer acquisition times. For example 3 accumulations of 10 seconds is similar to a single 30-second acquisition. Although it may take longer, and the SNR is slightly worse it can be of benefit when using samples that are more prone to laser damage (Table 3.3). It can be seen in Figure 3.6 that although the SNR is very low overall, it improved as the number of accumulations of 10-second acquisitions increased due to the reduction of random noise. This can be seen by the smoothness of the peaks in the spectra for 2, 5 and 10 accumulations.

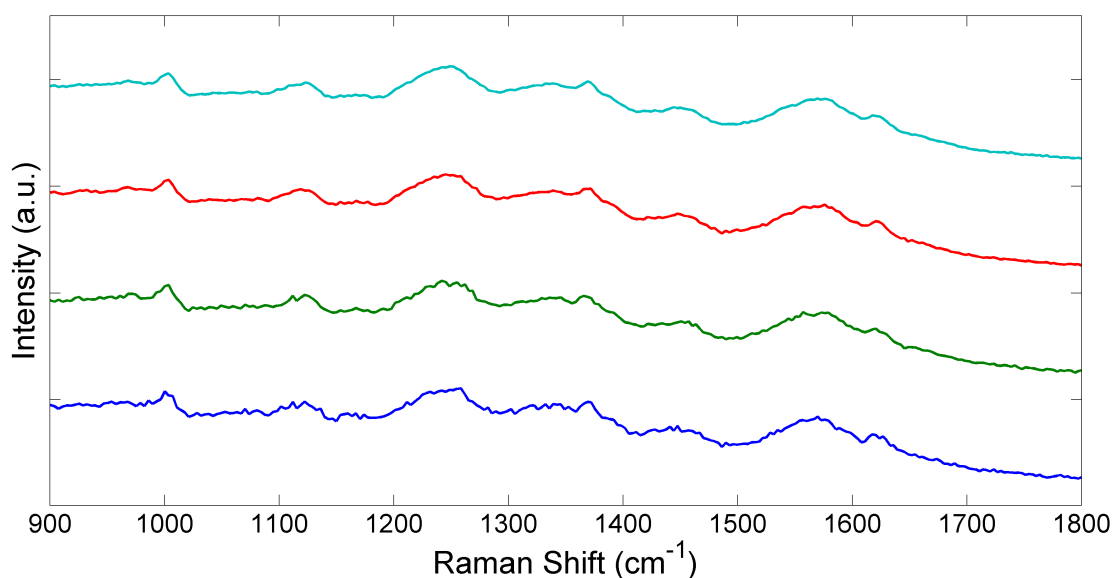


Figure 3.6. Mean Raman Spectra of Lymph Node Tissue Sections To Identify the Effect of Increasing the Number of Accumulations on Quality of Spectra; 1 (blue), 2 (green), 5 (red) and 10 (turquoise) accumulations

3.1.3. Raman System Calibration

The Raman spectrometer requires calibration at the beginning of each day. Calibrating the system allows a comparison to be made between measurements in order to avoid any deviation over time, this is particularly important in medical applications. Should there be any change in accuracy or precision this can be accounted for and adjusted at the time. If left these deviations could be mistaken for clinical differences in the spectra rather than technical. The calibration standards used have known characteristic peaks and are listed below with Figure 3.7 displaying some examples of the spectra obtained:

- Silicon, which has a very strong peak at 520.4 cm^{-1} , is measured to identify any shift in the spectra. Should there be any shift from this peak the system is recalibrated in order to centre the silicon peak at $520.4\pm 1\text{ cm}^{-1}$ (Figure 3.7A).
- Cyclohexane also has precise characteristic peaks, which are distributed across the spectral range. It is measured daily to identify any drifts in the spectra (Figure 3.7B).
- A neon-argon lamp has narrow emission lines at precise wavelengths, which are equally distributed across the spectral range, which again is used as a check over time to identify if the system has drifted (Figure 3.7C).
- Plastic pipette tip is also measured daily as it has a characteristic peaks distributed across the spectral range (Figure 3.7E).
- Green glass is measured to identify changes in intensity response of the system (Figure 3.7D). (Kendall, 2002 & Hutchings, 2009)

As the system was calibrated to the silicon peak on a daily basis the other spectra measured from the standards have been consistent across the three years of analysis.

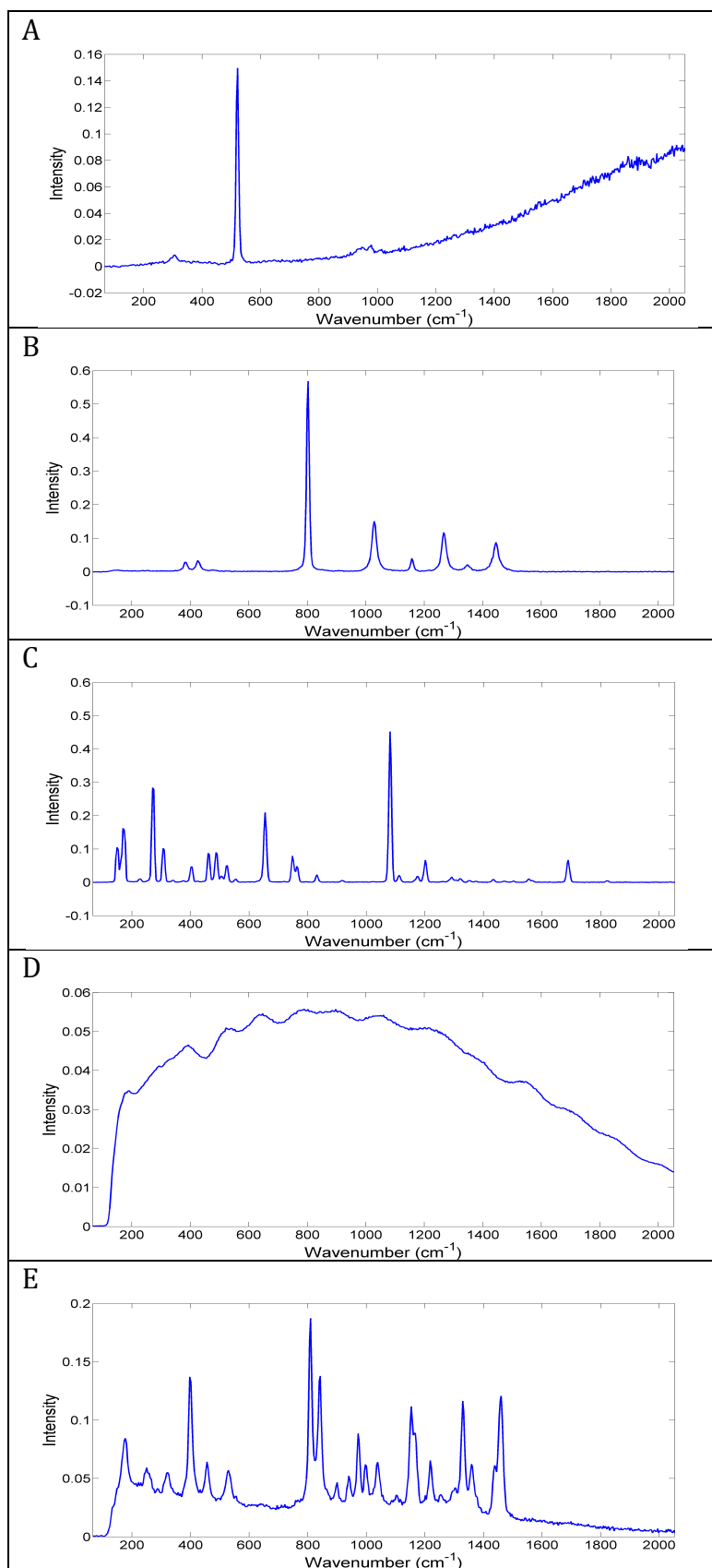


Figure 3.7. Spectra of Raman Calibration Standards; A) silicon, B) cyclohexane, C) neon-argon, D) green glass, E) pipette tip

3.2. Evaluating Drop Coating Deposition Spectroscopy (DCDS) for Microanalysis of Protein Solutions

3.2.1. Sample Preparation

Albumin, fibrinogen and γ -globulin are the three most abundant proteins found in human blood (Martini, 2006). Albumin, from human serum (A9511), fibrinogen from human plasma (F3879) and γ -globulin from human blood (G4386) were obtained from Sigma-Aldrich, UK. Albumin and γ -globulin were dissolved in analytical grade water (Pestanal water 34478, Fluka Analytical, Germany) to a concentration most comparable to their concentration in blood (Table 3.4). Fibrinogen was dissolved in 0.9 % saline (Fannin Ltd, Ireland) as it does not dissolve in water (Sigma-Aldrich, 2013a). These solutions were then diluted further to explore the spectra at very low concentrations (Table 3.5). Using a calibrated micropipette, the DCDS method was utilised to deposit 1.5 μ l droplets of each solution onto CaF₂ (Crystran, UK) and low-e slides (Kevley Technologies, USA), to allow the samples to be measured in IR transmittance, IR reflectance and Raman. Slides were left on the bench for sixty minutes at 22 °C in an air condition controlled laboratory.

Table 3.4. Protein Concentrations in Human Blood

Blood Protein	Concentration in Blood
Albumin	~50 mg/ml
Globulins	~30 mg/ml
Fibrinogen	~3 mg/ml
Peptide Hormones & Glycoprotein's	<1 mg/ml

Source: Martini, 2006

Table 3.5. Protein Dilutions

Blood Protein	Dilution	Protein Concentration (mg/ml)
<i>Albumin</i>	A	50
	A1	25
	A2	12.5
<i>Fibrinogen</i>	F	2.5
	F1	1.75
	F2	0.88
<i>γ-Globulin</i>	G	50
	G1	25
	G2	12.5

3.2.2. Raman, Infrared Transmittance and Infrared Reflectance Spectroscopy

Line maps of the samples were acquired, where each pixel of the map represents one spectrum. Infrared line maps were collected using the Perkin Elmer Spectrum One FTIR spectrometer coupled to a Spotlight 400 molecular imaging microscope described in section 3.1.1. White light images were created using a charge-couple device (CCD) camera, located in the microscope. For both reflectance and transmittance IR spectroscopy, 16 scans per 6.25 μm pixel were obtained at a resolution of 4 cm^{-1} , and 2.2 cm^{-1} interferometer speed.

Raman line maps were produced using a customised Renishaw Streamline System 1000 micro-spectrometer as described in section 3.2.2. A x50 objective lens was used to create line maps with 10 μm step size with each spectra acquired for thirty seconds.

3.2.3. Identifying the Position and Absorbance of the Amide I Peak

The amide I peak position and absorption across the samples were analysed and identified using an in-house line profile analysis tool for Matlab. The peak position was determined by selecting the wavenumber of maximum absorbance in the 1600-1700 cm^{-1} range. Colour maps of the spectral data were produced to

show the shift in absorption and position of the amide I peak across the ring from the outer edge, through to the other edge of the ring (Figure 3.8).

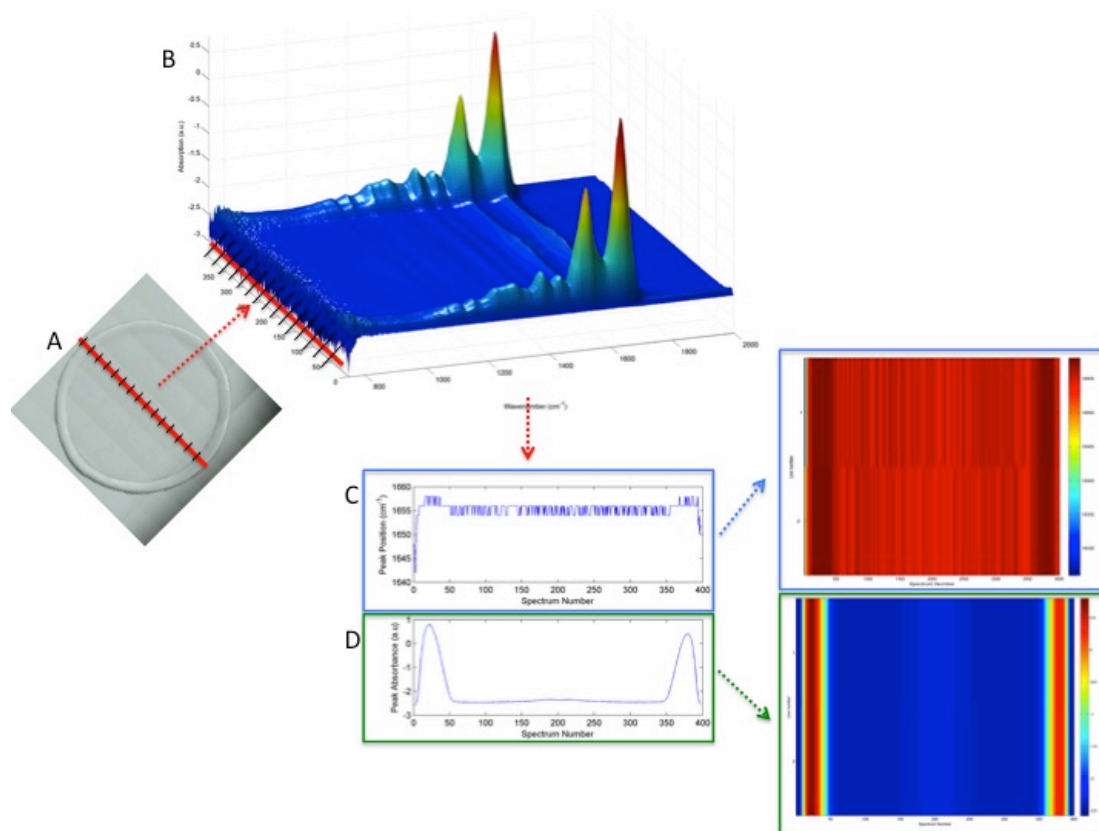


Figure 3.8. Schematic Diagram Showing the Stages of Line Profile Analysis.

A line map across the ring from edge to edge was obtained in IR and Raman (A), the spectra at each point along the ring were taken (B) and the position (C) and absorbance (D) of the amide I peak was analysed

3.3. Determining the Optimal Ring Characteristics for DCDS Using a Series of Albumin Solutions

3.3.1. Sample Preparation

Following on from these protein studies, albumin from human serum was diluted in analytical grade water to create a series of concentrations (Table 3.6). 1.5 μl of each solution was micro-pipetted onto a CaF_2 slide for IR transmittance and Raman spectral measurements, and a low-e slide for IR reflectance spectral

measurement. Slides were left on the bench for sixty minutes at 22 °C in an air-conditioned laboratory.

Table 3.6. Albumin Dilutions

Solution	Albumin Concentration (mg/ml)
A7	50.00
A6	25.00
A5	12.50
A4	6.25
A3	3.13
A2	1.56
A1	0.78
A0	0.00

3.3.2. Infrared and Raman Spectral Line Maps

Infrared and Raman line maps of the samples were acquired as described in section 3.2.2.

The white light images were used to determine the diameter of each ring. A white light image was taken of the United States Air Force (USAF) 1951 resolution test slide and the number of pixels across one bar was counted using Fiji ImageJ software (Schindelin *et al*, 2012). The width of that bar in microns was looked up on the T-20 USAF 1951 Chart and the number of microns per pixel in the USAF 1951 bar chart white light image was calculated to be 3.4 $\mu\text{m}/\text{pixel}$ (Applied Image Inc., 2013). This was then used to calculate the number of microns per pixel in the white light images of the samples, which was 2.6 $\mu\text{m}/\text{pixel}$. The number of pixels across each ring was again measured using Fiji ImageJ software and multiplied by 2.6 to determine the diameter of the ring in μm . This was calculated at several points and an average taken. From here on the diameter is referred to as the length across the entire droplet from outer edge to outer edge whereas the width is the distance from the inside to the outside edge of the ring.

3.3.3 Amide I Peak Analysis

The position and intensity of the amide I peak was measured across the diameter of each ring. Any peak shifts or scattering effects that may occur at the edges of the rings were then identified. The average amide I peak position for each of the seven albumin rings (A1-A7) was plotted to identify the most suitable concentration for each technique. The optimum area for spectral acquisition to avoid scattering and hence shifts in the peak position was then calculated.

3.4. Comparing DCDS and Cytospin Centrifugation as Blood Sample Preparation Methods in the Diagnosis of Leukaemia Using FTIR and Raman Spectroscopy

3.4.1. Ethics

Ethical approval was obtained from the South West 5 Research Ethics Committee (REC) (formerly Frenchay REC) for the analysis of an unspecified number of post-analysed samples of leukaemia whole blood, providing they were anonymously supplied. Approval was also granted for fifty samples of healthy whole blood from volunteers, provided that they consented to take part in the study after reading and fully understanding a patient information sheet (PIS) outlining the reasons for the study, the methods used and any risks that may be involved. The REC reference number for ethical approval is **06/Q2005/120** (Appendix C).

3.4.2. Sample Collection

Anonymous, post-analysed leukaemia blood samples were obtained from Cheltenham General Hospital; their pathologies identified and kept on a password-protected database by associated haematologists. All samples were obtained from a new patient clinic so that samples involved were less likely to be influenced by factors, such as chemotherapy, that would alter the biochemical fingerprints of the cancers. Samples were supplied in standard 5ml ethylenediaminetetracetic acid (EDTA) blood collection tubes to ensure the

samples did not clot. Simultaneously, whole blood from healthy volunteers, who consented to taking part in the study, were also collected in standard 5 ml EDTA blood collection tubes by phlebotomists. Samples were then screened by the associated haematologists to ensure their suitability in the study.

3.4.3. Sample Preparation

The protocol (Appendix D) was adapted and developed from the Sigma-Aldrich Histopaque®-1077 Product Information page (Sigma-Aldrich, 2011). Once obtained, all samples were maintained at 2-8 °C, prepared and fixed within six hours of collection and analysed within 36 hours of preparation. For each sample collected, 3 ml of whole blood was required for centrifugation. A further 150 µl was kept for cytospin centrifugation and remaining blood was diluted for Drop Coating Deposition Spectroscopy (DCDS) (see section 3.4.4.2).

3.4.4. Methods

3.4.4.1. Isolating Mononuclear Cells from Anti-Coagulated Whole Blood by Density Gradient Centrifugation

Whole blood was collected and within six hours, 3 ml was micropipetted into a 20 ml Steralin tube containing 5 ml of phosphate buffered saline (PBS) (P4417, Sigma Aldrich, UK), and mixed well by inversion. The 8ml blood-PBS mixture was layered onto 3 ml of Histopaque®-1077 (Sigma Aldrich, UK) in a 15 ml conical centrifuge tube held at a 45° angle to prevent the layers merging (Figure 3.9 part A). The solution was then centrifuged at 40 g (1300 rpm) for exactly thirty minutes. Following centrifugation, the upper layer (supernatant) was aspirated to within 0.5 cm of the opaque interface containing the mononuclear cells with a Pasteur pipette. 200 µl of the upper layer (containing plasma and PBS) was kept for DCDS and the rest discarded (Figure 3.9 part B).

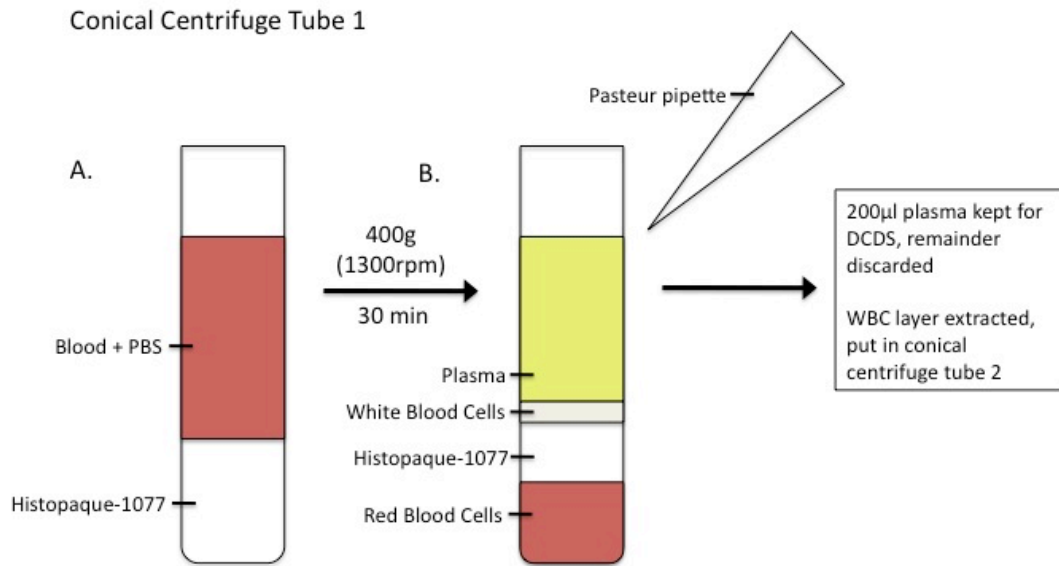


Figure 3.9. Conical Tube 1

With a Pasteur pipette, the opaque white blood cell layer was transferred into a second 15 ml centrifuge tube containing 10 ml PBS and mixed by inversion (Figure 3.10 part A). The new solution was then centrifuged at 250 g (1000 rpm) for exactly ten minutes. After centrifugation, the supernatant was again aspirated and discarded (Figure 3.10 part B). The lymphocyte pellet was re-suspended with 5 ml PBS and centrifuged a second time at 250 g (1000 rpm) for exactly ten minutes. This stage was repeated for a third time with 5 ml PBS to ensure removal of all other blood fractions from the white blood cell layer.

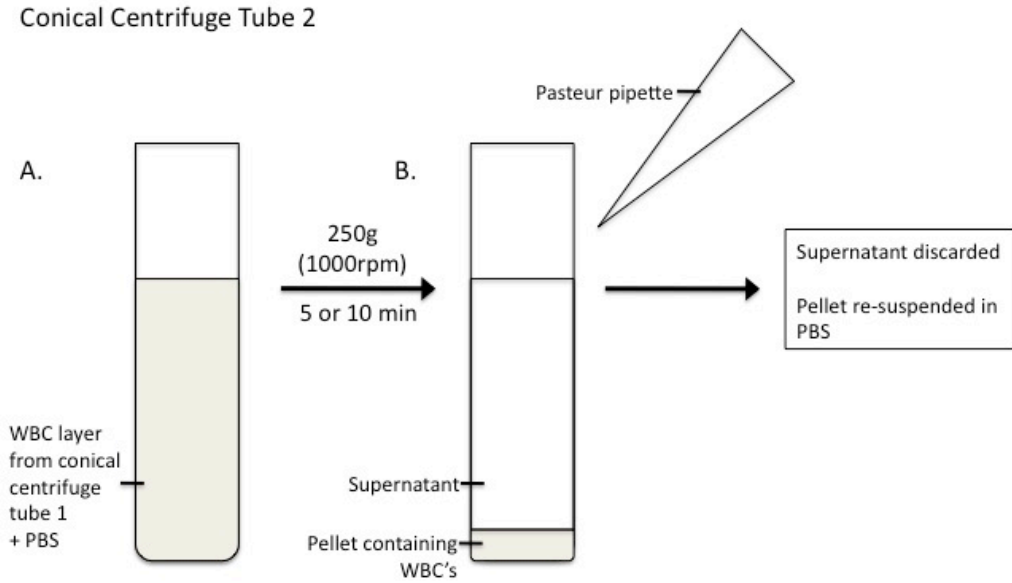


Figure 3.10. Conical Tube 2

After the final centrifugation the supernatant was discarded and the pellet re-suspended in 0.5 ml PBS. The white blood cell fraction was appropriately diluted with PBS in order to acquire a monolayer on the CaF₂ slides. 150 µl of buffy coat solution and 150 µl of the original whole blood were placed into cytopspin mounts and cytopspun for 2 mins at 500 rpm (Figure 3.11). Slides were left on the bench for a minimum of two hours at 22 °C in an air-conditioned laboratory.

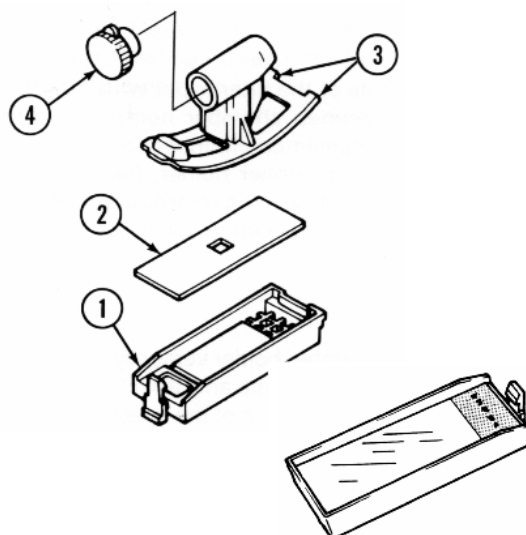


Figure 3.11. The Cytospin Mounts; CaF₂ slide inserted into base holder (1), filter paper (2) lies on top of slide to absorb excess fluid, fluid chamber (3) clips on top and cap (4) is tightened to hold sample securely

3.4.4.2. Drop Coating Deposition Spectroscopy

100 %, 50 % and 25 % dilutions of whole blood and plasma were made and 1.5 μ l droplets of these solutions were deposited onto CaF₂ slides with a calibrated micropipette. The droplets were left to dry as the solutes were pushed to the outer edge in order to provide a more concentrated sample for spectroscopic analysis. Slides were left for sixty minutes at 22 °C in an air condition controlled laboratory.

3.4.4.3. Spectroscopic Analysis

Point spectra of the samples were acquired using the Perkin Elmer Spectrum One FTIR spectrometer coupled to a Spotlight 400 molecular imaging microscope described in section 3.1.1. Images were created under white light using a charge-couple device camera, located in the microscope. Thirty point spectra were taken from each sample type with 16 scans per 6.25 μ m pixel obtained at a resolution of 4 cm^{-1} , and 2.2 cm interferometer speed. Background measurements of the CaF₂ slides were taken and ratioed against each point

spectra in order to minimise any interference. Point spectra were also acquired from the customised Raman Renishaw Streamline System 1000 micro-spectrometer as described in section 3.1.2. A x50 objective lens was used to obtain thirty point spectra from each sample for thirty seconds.

3.4.4.4. Data Analysis

The data was subject to both univariate and multivariate statistical analysis using Matlab in order to generate a classification model. Spectral pre-processing, as described in section 2.3.1, involved normalisation and mean centring of the data in order to remove any differences due to absorbance and intensity.

Peak ratio analysis was carried out using an in house Matlab tool in order to remove any weak or saturated IR data. Ratios of the amide I and amide II peaks were calculated and plotted in histograms for each sample type. The data was then compared to various sources and a decision made to exclude saturated data based on the individual groups.

Univariate analysis of the data was carried out using Matlab. Mean spectra were produced from an average of the point spectra for each pathology and each sample type within that pathology. Difference spectra comparing the data between the pathologies for every sample type were also produced. Finally the peaks in the spectra were assigned using an in house Matlab tool.

Multivariate analysis involved principal component analysis (PCA) followed by linear discriminant analysis (LDA) and leave one sample out cross validation (LOSOCV) as described in section 2.3.3. Twenty principal components (PC) were calculated and their scores and loads plotted to identify any natural separation in the data and identify the biochemical causes of the separation. The PCs were then used for LDA where a linear function was used to maximise the variances in the data between the different groups whilst minimising the variance in the data within the groups. Finally LOSOCV was used to calculate the sensitivity (SN) and specificity (SP) of the classification model.

3.5. Lymphoma Diagnosis Using FTIR and Raman Spectroscopy

3.5.1. Ethics

An application was also made to the South West 5 Research Ethics Committee (REC) (formerly Frenchay REC) for the analysis of an unspecified number of post-analysed samples of lymphoma fine needle aspirate biopsies (FNAB), providing they were anonymously supplied. Ethical approval was also obtained for fifty lymph node tissue biopsy samples providing consent was obtained from the volunteer patients involved. The REC reference number for ethical approval is also **06/Q2005/120** (Appendix C).

3.5.2. Sample Collection & Preparation

Anonymous, FNABs were obtained from Cheltenham General Hospital and placed in PBS. No lymph node tissue biopsies were obtained.

3.5.3. Methods

3.5.3.1. Cytospin Centrifugation

Once samples were obtained, they were subject to the cytopspin centrifugation method utilised for the leukaemia blood samples. 150 μ l of the aspirate solutions were placed into cytopspin mounts and cytopspun for 2 mins at 500 rpm. The slides were left to dry on the bench for a minimum of two hours prior to spectroscopic analysis.

3.5.3.2. Drop Coating Deposition Spectroscopy

DCDS was experimented with on one of the FNABs. The method used was as described in section 3.4.4.2.

3.4.5.2. Spectroscopic Analysis

As described in section 3.4.4.3.

4. RESULTS AND DISCUSSION: BLOOD PROTEIN ANALYSIS

In preparation for the analysis of leukaemia blood samples, drop coating deposition spectroscopy (DCDS) has been explored using both Raman and Fourier Transform Infrared (FTIR) spectroscopy to compare the limitations of the techniques for biological analysis of samples with low protein concentration. The first section, 4.1, explores the results from the initial study involving the three blood proteins; albumin, fibrinogen and γ -globulin. The main focus being on the variation of the amide I peak across the samples. Any shifts in the amide I peak position will determine whether scattering has occurred, whilst differences in absorption of the amide I peak will identify whether the sample deposits within a 'coffee-ring' pattern. Scattering may cause peaks to shift, which can be problematic as this may be mistaken for a change in the biochemistry due to disease. The second section, 4.2, explores the DCDS method further with albumin solutions only. The aim of this study was to identify what effect sample concentration has on droplet formation and spectral data with the aim of reducing sample volume whilst maintaining sensitivity of the techniques. By comparing the data obtained from FTIR and Raman it is hoped that the DCDS method can be used across complementary platforms whilst achieving consistent results.

4.1. Evaluating Drop Coating Deposition Spectroscopy (DCDS) for Microanalysis of Protein Solutions

4.1.1. Sample Preparation Using DCDS

The DCDS method, as described in section 3.2.1, works by capillary flow, a process of sample drying previously described by Deegan *et al* (1997). Once the droplets have been micropipetted onto a slide, they dry from the outside in creating a flow of solutes to the outer edge, which then becomes highly concentrated compared to the solution in the centre. Once the sample has completely dried, the solutes should theoretically be located in a 'coffee-ring'

pattern on the slide, thus making spectroscopic analysis easier due to the concentration increase. The effect that occurs can be seen in the white light images in Figures 4.1-4.3.

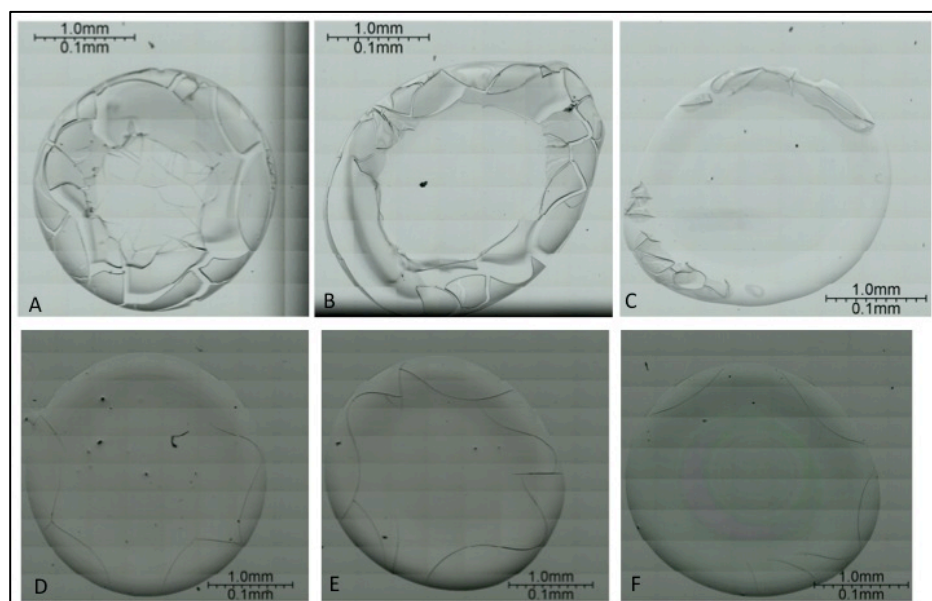


Figure 4.1. White Light Images of Albumin Droplets; Images A-C represent albumin droplets on CaF₂ slide used for Infrared (IR) transmittance and Raman, images D-F represent the low-e slide used for IR reflectance. A) & D) dilution A, B) & E) dilution A1, C) & F) dilution A2 as described in Table 3.3.

The images in Figure 4.1 represent the albumin droplets, with dilution A having a concentration of 50 mg/ml, A1 25 mg/ml and A2 12.5 mg/ml. The droplets on the CaF₂ slide appear to improve in structure as the concentration decreases. This suggests that the protein concentration in the ring is too high for the amount of fluid and thus the droplets have cracked as they dried. The widths of the rings are also very large and uneven in the more highly concentrated samples. The droplets on the low-e slides have a significantly better formation. The droplets have less cracks and are more spherical and evenly distributed across all concentrations. The droplets are larger, most likely due to the hydrophilic nature of the low-e slides (as opposed to the hydrophobic CaF₂ slides) causing the fluid to spread further. This may also explain why the droplets have a better structure.

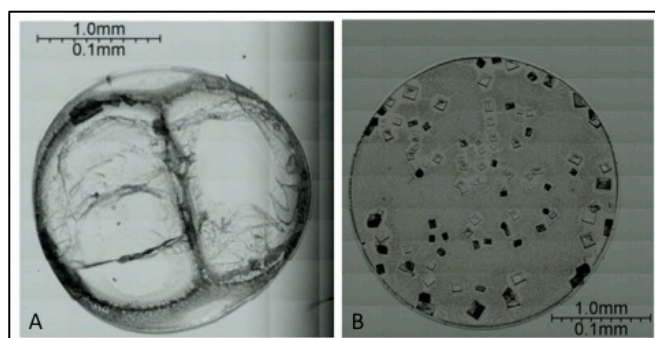


Figure 4.2. White Light Images of Fibrinogen Droplets; A) Dilution F (2.5 mg/ml) deposited onto CaF₂ slide used for IR transmittance and Raman, B) dilution F2 (0.88 mg/ml) deposited onto low-e slide used for IR reflectance.

In Figure 4.2A it can be seen that ring F on the CaF₂ slide has not formed properly due to an error in the deposition method. When dispensing the fluid onto the slide a pocket of air became trapped in the liquid causing the droplet to dry unevenly around an air bubble. This was avoided when pipetting other droplets. However the other fibrinogen droplets have a very different structure to the albumin droplets. This is most likely due to the use of saline as a solvent rather than water. It has resulted in crystalline structures and ferning patterns forming in the droplets and very small ring widths as represented by the fibrinogen droplet in Figure 4.2B. Ferning patterns like this have been seen in other studies involving tear droplets and have been used to aid in the diagnosis of ocular disease (Pearce & Tomlinson, 2000 & Filik & Stone, 2007). The crystals may contain areas of higher fibrinogen concentration and thus may be useful in spectroscopic analysis. As with the albumin solutions, the droplets were larger on the low-e slides with the crystalline structures also appear to be smaller and more dispersed throughout these droplets.

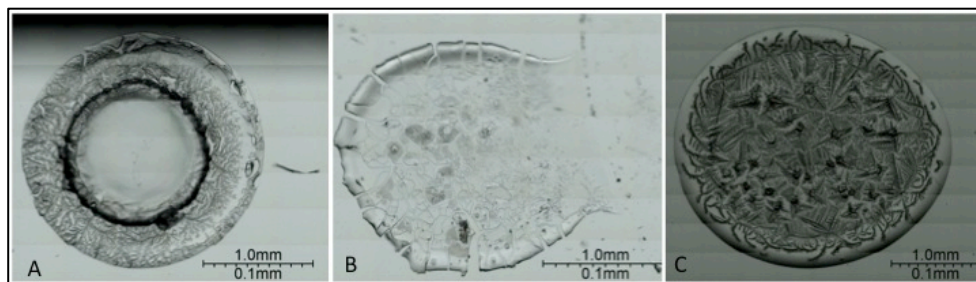


Figure 4.3. White Light Images of γ -Globulin Droplets; A) Dilution G (50 mg/ml) deposited onto CaF_2 slide, B) dilution G2 (12.5 mg/ml) deposited onto CaF_2 slide, C) dilution G1 (25 mg/ml) deposited onto low-e slide used for IR reflectance

The droplets on the CaF_2 slide are all very different to one another identifying a problem with the method (Figure 4.3 A-B). Looking at the product information sheet supplied by Sigma-Aldrich it would appear that its solubility is pH-sensitive, and is soluble in 0.9 % sodium chloride (Sigma-Aldrich, 2013b). As water was used the γ -globulin has not dissolved properly resulting in the formation of highly varied droplets with crystalline structures observed in some rings, ferning in others and a mixture of ring widths throughout. This was not repeated due to the problems using saline for the fibrinogen droplets. As with the fibrinogen droplets ferns were seen on the low-e slides (Figure 4.3C).

Upon visual inspection it appears that overall not all of the rings are perfectly spherical. From repeated pipetting it has become evident that for the droplets to be spherical the pipette needs to be held at a 90° angle from the slide to ensure even deposition, thus the nature of these droplet formations is due to the errors in the way the droplet was deposited. This is a technique that once mastered is very easy to repeat, as can be seen in section 4.2.1 (Filik & Stone, 2007).

Across all of the samples there is a noticeable difference in the droplet size between the CaF_2 and low-e slides. The droplets appear larger on the low-e slides, probably due to the surface tension and the way in which the samples adhered to the slides. This property needs to be addressed as according to Zhang *et al* (2003), the substrate on which the drops are deposited is important for the

method to be effective, and commonly gold or stainless steel coated in Teflon is used. However, these types of slides are expensive and according to Filik & Stone (2007), CaF₂ and low-e slides are still capable of producing good quality spectra. Not only is the slide important in droplet formation but the solvent in which the protein is dissolved in is too. If dilution of the blood samples is required then the media in which this is done must not interfere with the properties of the blood nor with the flow of the liquid on the slides.

4.1.2. Raman, Infrared Transmittance and Infrared Reflectance Spectroscopy

Using the parameters described in section 3.2.2, spectra were obtained from Raman, IR transmittance and IR reflectance in order to explore drop coating deposition spectroscopy (DCDS) as a viable method for blood sample analysis. It is already known that measuring the buffy coat (the leucocyte fraction of blood) by infrared spectroscopy, chronic lymphoblastic leukaemia (CLL) can be diagnosed to a high sensitivity, with 99.5 % of samples correctly classified (Babrah, 2009). However, the process of separating out the different fractions of blood is time consuming and if DCDS can generate the same quality of results it could be used as an alternative as it is quick and easily repeatable. In chapter 5 both of these methods are explored with blood samples to diagnose leukaemia.

In using this method, the interference of Mie scattering needs to be addressed. This dispersion artefact occurs when the wavelength of incident light is equal to the diameter of a particle, or object of interest. When light collides with the object it is scattered away from the detector causing a decrease in signal intensity and an artificial shift in the spectra, noticeably around the amide I peak. Any change to this peak in the past was thought to be due to a change in protein conformation (Bassan *et al*, 2010). If this scattering were to arise and not be accounted for in the leukaemia blood samples, the spectra could be misinterpreted, potentially leading to a misdiagnosis of cancer. Shultz *et al*, (1996) compared the spectra of normal and CLL cells using FTIR. They identified a shift in the amide I peak from 1656 cm⁻¹ in normal lymphocytes to 1652 cm⁻¹ in

CLL cells. Had there been a shift in the spectra due to Mie scattering, the diagnosis may have been falsely made, or not made at all. Thus in this study we have concentrated the analysis of scattering to this region of the spectra.

Mie scattering is most commonly observed in IR reflectance spectra due to a change in the trajectory of the radiation caused by differences in the refractive index between a sample and its surrounding medium, as well as the increased intensity observed when radiation passes through a sample and back (Romeo *et al*, 2008, Bassan *et al*, 2010). A small amount of scattering may also be observed in IR transmittance while no affect should be seen in the Raman spectra. Thus IR-reflectance spectra are measured as a standard to aid in the observation of scattering in Raman and IR transmittance spectra, which if discovered can later be corrected for. Figure 4.4 shows some typical mean spectra from each of the three methods. Ring A1, with a concentration of 25 mg/ml was used. There is a split and a shift in the amide I peak in the IR reflectance spectrum, which can be seen at 1657 cm^{-1} in the Raman and IR transmittance spectra demonstrating the effect that scattering has on the results.

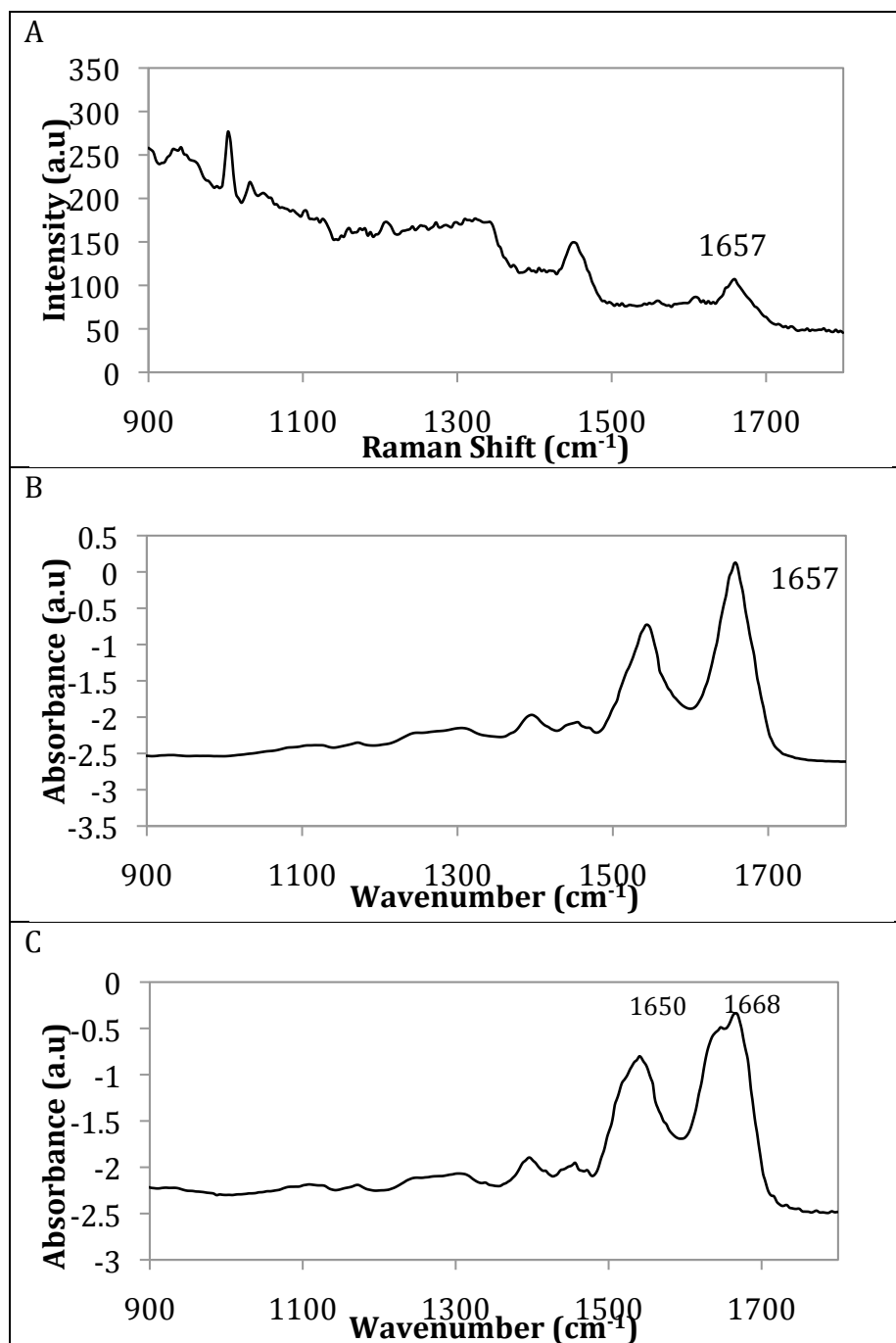


Figure 4.4. Typical Spectra from Albumin Ring A1; A) Raman, B) IR transmittance, C) IR reflectance

4.1.3. Identifying the Position and Absorbance of the Amide I Peak

As described in section 3.2.3, peak absorption and peak position colour plots were created from the line maps and analysed to visually inspect the deposition of the albumin across the droplet. In order to explain how they were used, the plots from albumin droplet A2 measured in transmittance are displayed in Figure 4.5. In part A, the dark red areas indicate high spectral absorption in the 1600-1700 cm^{-1} region, which appears to decrease at the edges of the ring. The blue regions indicate very little or no absorbance suggesting that the albumin was only deposited within the ring. In part B, the plot is more difficult to analyse. The amide I peak position within the ring appears to reside at 1657 cm^{-1} , however throughout the plot there appears to be a peak at approximately 1655 cm^{-1} even though there is no albumin deposited here. The way in which the plots were constructed meant that the peak with the highest absorption within the 1600-1700 cm^{-1} region was selected across the entire droplet, regardless of whether it was from background noise. These plots were mainly used to identify that the deposition of the proteins had successfully been achieved.

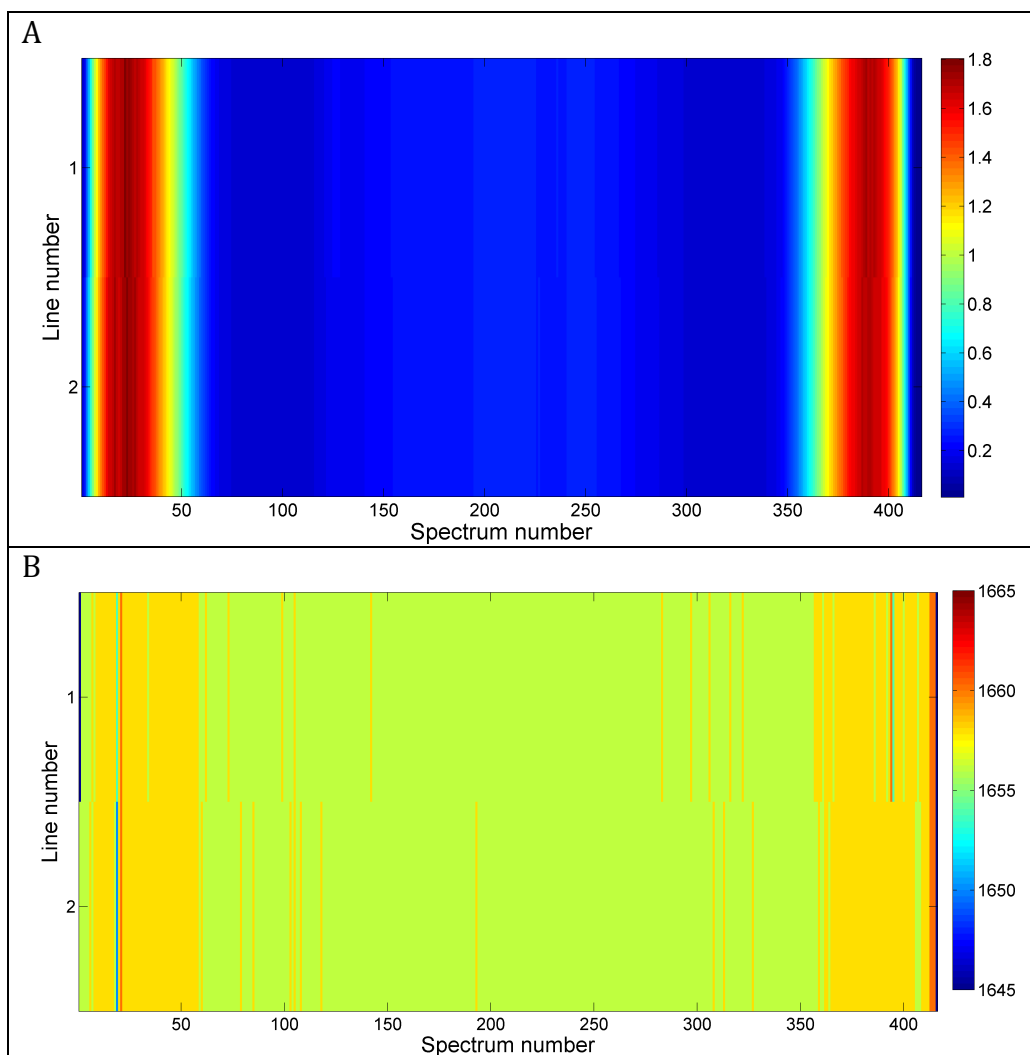


Figure 4.5. Peak Absorption and Position Plots of the Amide I Peak for Albumin Droplet A2 Measured in IR Transmittance; A) peak absorption, B) peak position

Although it can be seen by these plots that the amide I peak and absorbance was consistent within the ring and that no albumin resided outside of the ring the plots show a lot of information that can more easily be summarised by line profiles. Line profiles were a plot of the numerical values and thus were analysed to accurately identify the position and absorbance of the amide I peak. These were then compared to literature (Appendix B) as well as data obtained from measuring these proteins across the Raman and FTIR systems (Appendix E). Line profiles for albumin measured in IR transmittance are shown in Figure 4.6.

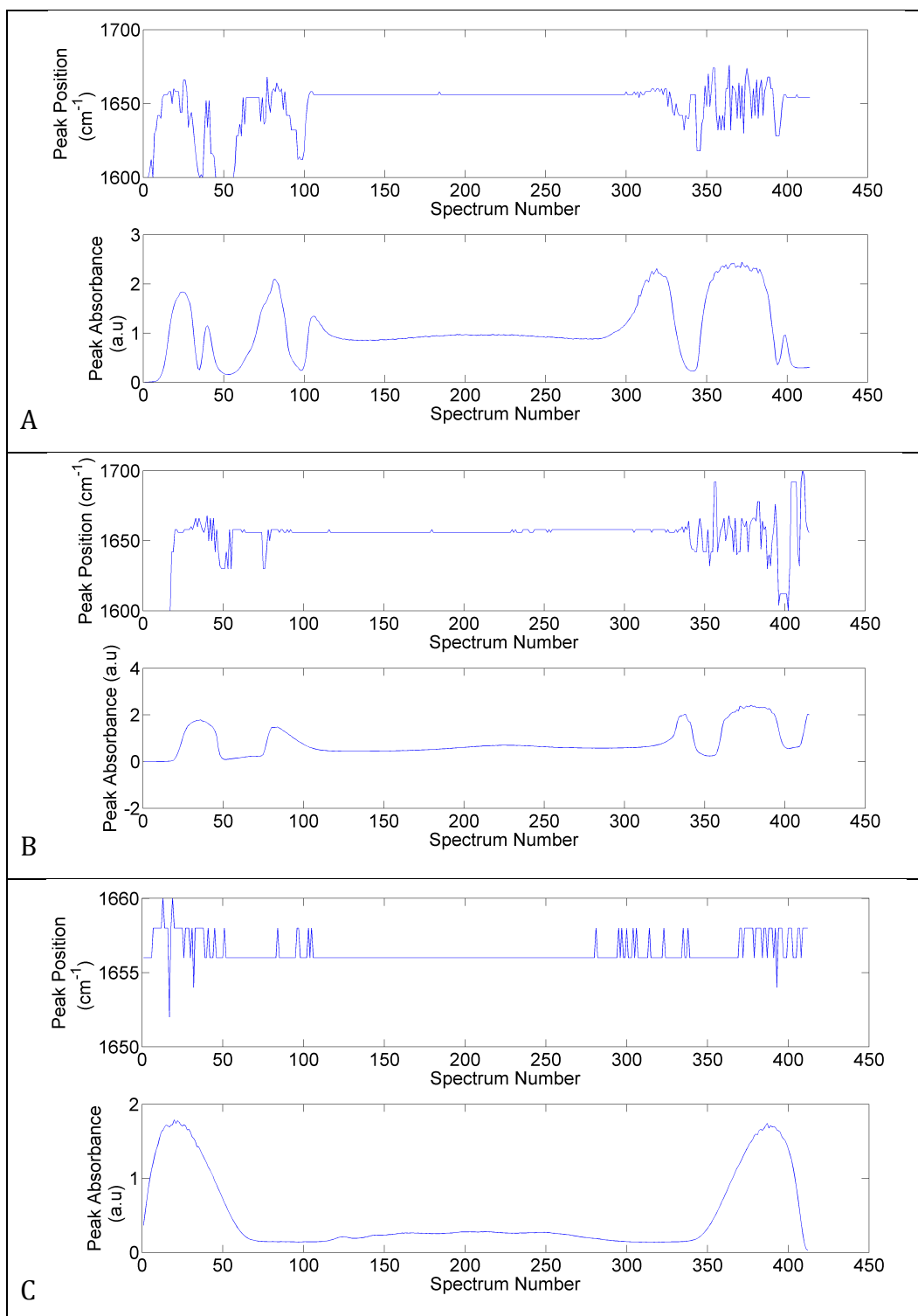


Figure 4.6. Line Profile Plots of Albumin Rings Measured in IR Transmittance; showing the peak position and absorption of the amide I peak for A) ring A, B) ring A1 and C) ring A2.

The position and absorbance of the amide I peak for ring A varies significantly throughout. The areas of low absorbance coincide with cracks in the ring that were formed during the drying process (Figure 4.1) causing the peak position to shift randomly in these areas between 1600-1680 cm^{-1} . It is more likely that this is due to background noise as opposed to the presence of Mie scattering. In areas where the absorbance is high, the peak position also varies slightly. However this is due to the sample being too thick for the spectrometer causing the peaks to flatten at the top when the maximum absorbance is reached. The centre of the peak is therefore more difficult to determine causing an apparent shift in the amide I peak position. A similar picture can be seen for ring A1, although the maximum absorbance is slightly lower. For ring A2 the position of the amide I peak varies only slightly between 1656 cm^{-1} and 1658 cm^{-1} . The FTIR spectrometer measures at intervals of two wavenumbers which explains the variation of the amide I peak position between these two points. There are however a couple of points within the ring where the amide I peak position has drifted outside of this range which can also be explained by spectra being taken in the cracks.

In order to further demonstrate the differences in absorbance and peak position between the albumin rings measured in IR transmittance, 3D spectral plots are displayed below (Figure 4.7). The spectrum number indicates each individual spectrum taken across the entire ring, i.e. 400 spectra were taken across the line map. In the 3D plot for ring A (Figure 4.7A) the top of the peaks at the higher wavenumbers can be seen to be flattened out, whereas in ring A2 (Figure 4.7B) where the concentration is lower the peaks are much more defined.

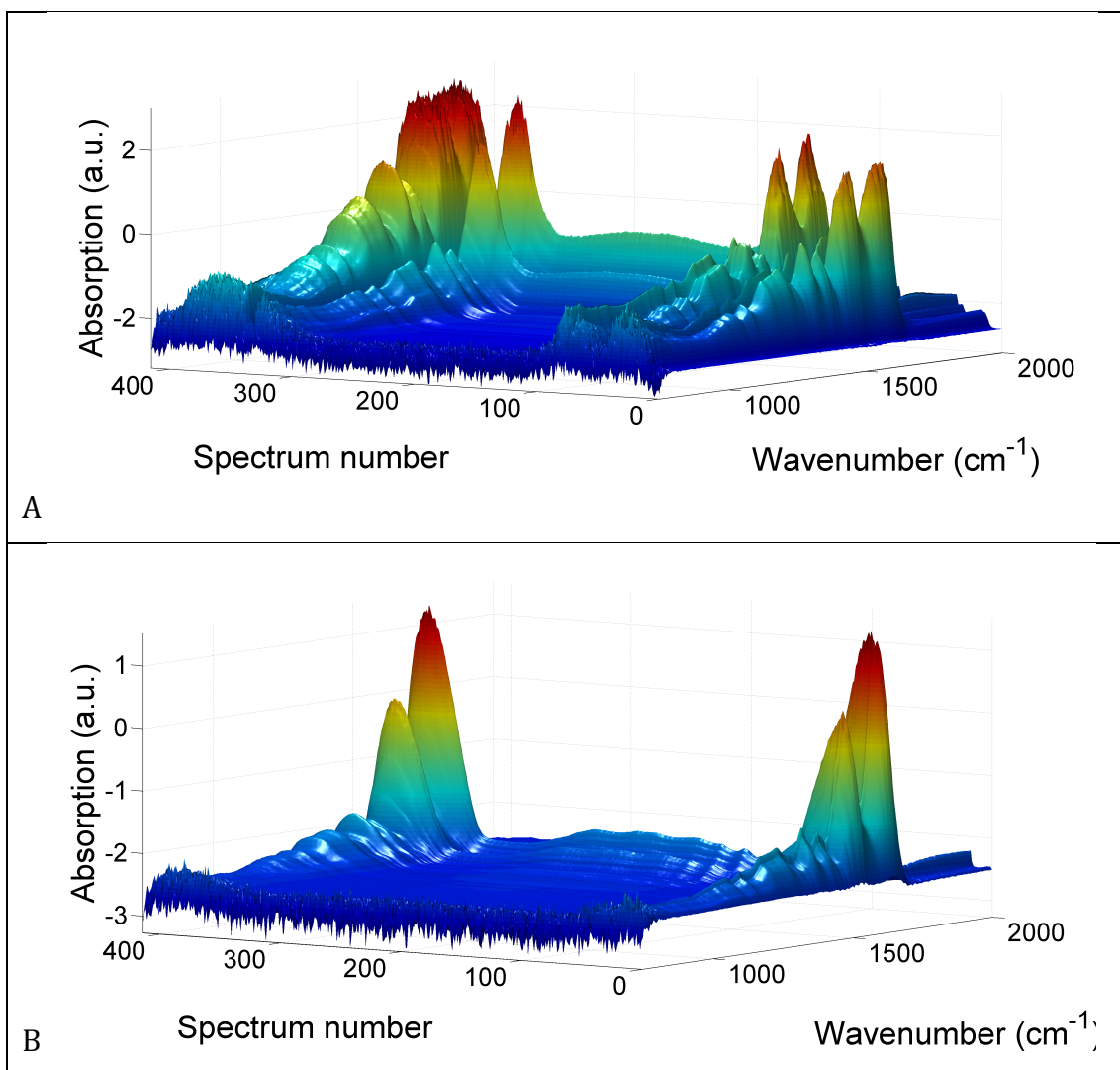


Figure 4.7. 3D Spectral Plots of Albumin Rings A and A2 Measured in IR Transmittance; A) ring A, B) ring A2

The line profile plots for the albumin rings measured by Raman spectroscopy can be seen in Figure 4.8. As with the IR transmittance data, the areas where there are cracks in the rings can be seen by variation in the intensity and position of the amide I peaks. Although there appears to be less variation or interference from cracks, this is just due to the line maps being measured in different places across the rings. The intensity of the peaks decreases as the concentration of albumin in the ring decreases suggesting that better quality Raman spectra can be obtained with higher concentrations of albumin, from 25 mg/ml up. As with the IR transmittance line profiles it is clear that albumin has

only resided at the edge of the droplets suggesting that the DCDS method was successful in evenly depositing the albumin within a ring.

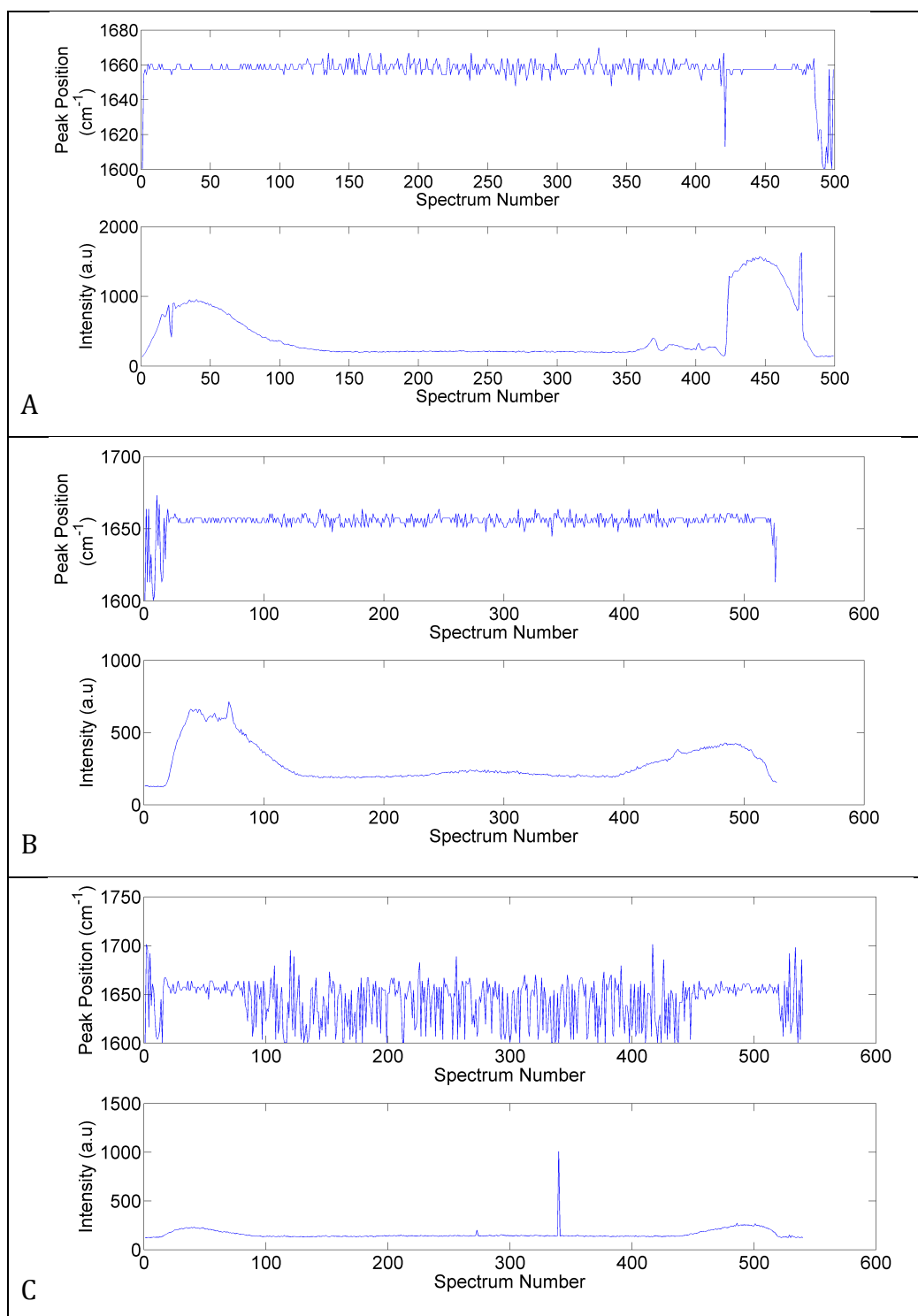


Figure 4.8. Line Profile Plots of Albumin Rings Measured in Raman; showing the peak position and intensity of the amide I peak for A) ring A, B) ring A1 and C) ring A2.

For completeness, and in order to compare the data between the three spectroscopic methods, the line profile for albumin ring A in IR reflectance is shown in Figure 4.9. It can be seen that within the ring the position of the amide I peak is much higher than in the other two methods, residing between 1680-1690 cm^{-1} . This consistent shift in the peaks is caused by scattering and confirms the lack of scattering present in both IR transmittance and Raman spectroscopy. Again spectra measured in cracks in the rings can be observed where the absorbance and peak position is low within the rings.

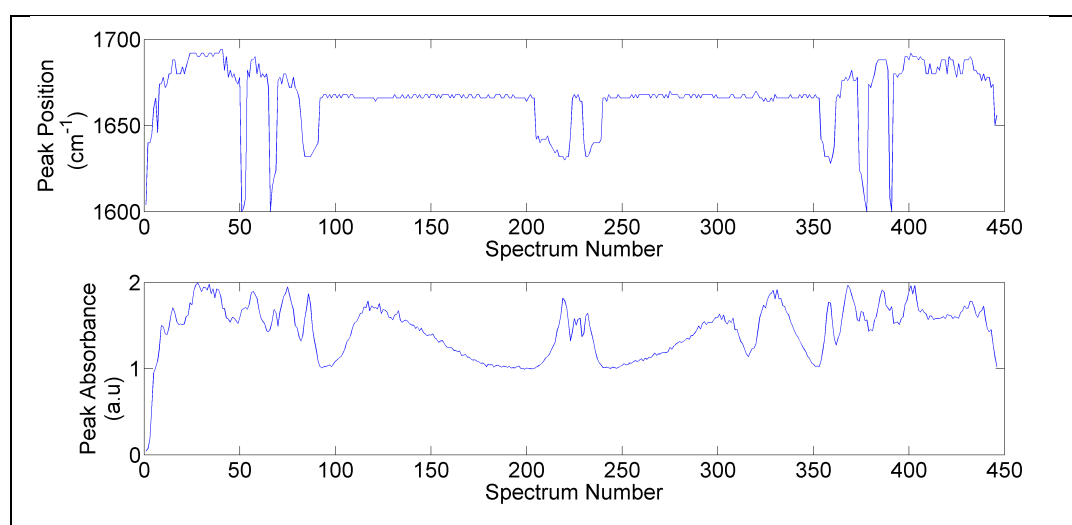


Figure 4.9. Line Profile Plot of Albumin Ring A Measured in IR Reflectance; showing the peak position and absorption of the amide I peak

From the line profiles, the range of values of the amide I peak position for all droplets measured in Raman, IR transmittance and IR reflectance were identified and recorded in Table 4.1.

Table 4.1. Amide I Peak Positions (cm⁻¹)

Protein	Socrates (2001)	Raman	IR transmittance	IR reflectance
<i>Albumin</i>	1645-55			
Pure Albumin		1657	1656-58	1656-58
A		1657	1656-58	1680-92
A1		1657	1656-58	1680-90
A2		1654-57	1656-58	1680-90
<i>Fibrinogen</i>	1660-70			
Pure Fibrinogen		1663-66	1664	1664
F		1663-66	1652	1658
F1		1663-66	1652	1664-6
F2		1666	1652	1658
<i>Globulin</i>	1665-1675			
Pure Globulin		1672-75	1672-74	1672-74
G		1672-75	1642	1688-94
G1		1672-75	1652	1688-94
G2		1672-75	1644	1688-92

According to Socrates (2001), the amide I peak for an α -helical protein such as albumin, should reside between 1645-1655 cm⁻¹ in both infrared and Raman. The Raman data showed that it actually resided at 1657 cm⁻¹, which although is slightly higher this is due to the instrument itself. The pure protein data confirms this as does the consistency of the data across the different concentrations suggesting that it is a technical difference rather than an error in the DCDS method (Appendix E). At the lowest concentration, A2, there appears to be some variation in the peak position. From the spectra it was clear that the droplet was too thin for the weak Raman signal and thus this variation may be due to some small degree of scattering from cracks in the droplet. The average peak position was also shown to be 1657 cm⁻¹ by the IR transmittance method, which is again slightly higher but does coincide with the pure albumin data. This is consistent across all three droplets. The IR reflectance profiles identified a shift in the amide I peak of 20 to 30 wavenumbers. This was expected due to the dispersive effects that take place when the incident radiation is equal to the change in

refractive index at the edge of the droplet. The fibrinogen and globulin profiles appear to further confirm the scattering in the reflectance data.

Fibrinogen is a random chain protein, which should have an amide I peak between 1660-1670 cm^{-1} , whereas globulin has a β -sheet structure and should have an amide I peak between 1665-1675 cm^{-1} (Socrates, 2001). The poor ring structure and the uneven distribution of the protein throughout the fibrinogen and globulin droplets make it difficult to conclude the accuracy of the DCDS method both physically and spectrally. The proteins did not deposit in a ring and the position and absorption of the amide I peak was highly varied due to scattering.

4.1.4. Summary

From the data obtained it has been identified that the DCDS method can be used for the microanalysis of proteins. Although the fibrinogen and globulin results were poor, it was determined that saline is not an appropriate solvent as it causes interference in the droplet formation. The albumin droplets showed consistent results across the dilutions in both Raman and IR transmittance. When utilising this method for blood analysis this will make the droplet deposition process easier and more accurate as it is very difficult to dilute every blood sample to achieve consistent droplet sizes and thickness due to the highly variable biochemical concentrations within each sample. At lower concentrations the droplets are too thin for Raman analysis, and at higher concentrations they are too thick for IR transmittance analysis. However by depositing a series of diluted droplets of each patient's blood on a slide and measuring the most suitable one for the spectroscopic method we can be confident that the same results will be achieved. To further understand the limitations of the droplet concentrations on the Raman and IR transmittance spectra, the method was explored further in section 4.2 using a larger range of albumin dilutions.

4.2. Determining the Optimal Ring Characteristics for DCDS Using a Series of Albumin Solutions

In order to study the protein droplets and their formation on the slides further a series of albumin dilutions covering a larger range of concentrations were analysed. The purpose of this was to identify the upper and lower limits of the DCDS technique using both Raman and FTIR with the future aim being to utilise the technique for biological analysis of samples with unknown or uncontrollable concentrations, such as blood.

4.2.1 Physical properties

The white light images show a severe loss of shape and structure as the concentration of albumin decreases (Figure 4.10). At the highest concentration, both the diameter of the droplet and the width of the ring are larger. At lower concentrations, more cracks are visible in the rings. The white light images were used to explore the relationship between ring width and albumin concentration, which in Figure 4.11 are shown to be directly proportional at the higher concentrations. At more dilute concentrations, rings A0 to A5, the rings have not properly formed suggesting that at concentrations below 25 mg/ml (below the normal concentration range of albumin in human blood) the technique is limited (Martini, 2006). Interestingly, a linear fit of the data indicated that at 0 mg/ml there should be a droplet with a ring width of 22.9 μm . This was tested with a 1.5 μm droplet of the same analytical grade water (Figure 4.10A). A droplet was seen and when measured was found to have a ring width of 18.2 μm . However as it was more difficult to measure the ring width accurately at lower concentrations due to poor structure and cracks, these results are less accurate and below 12.5 mg/ml the data is not linear.

Interestingly, the droplets are more spherical and consistent in shape here compared to the previous study. Although this did not appear to affect the spectra it does highlight how quickly the method can be mastered in order to obtain a reliable method of deposition.

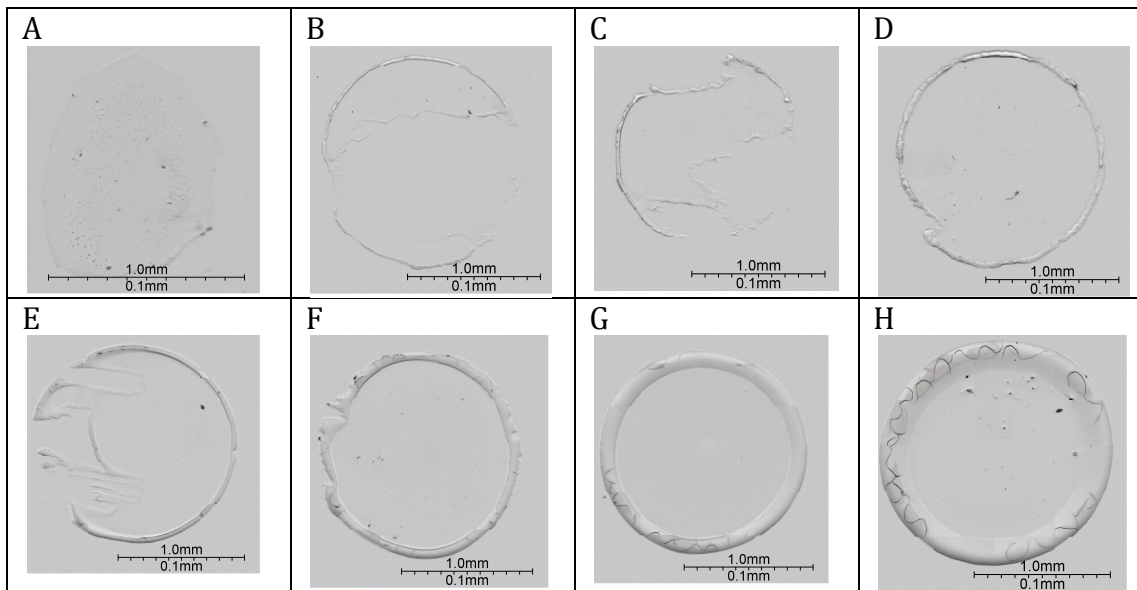


Figure 4.10. White Light Image of Albumin Droplets on CaF₂ Slide; A) ring A0, B) ring A1, C) ring A2, D) ring A3, E) ring A4, F) ring A5, G) ring A6, H) ring A7

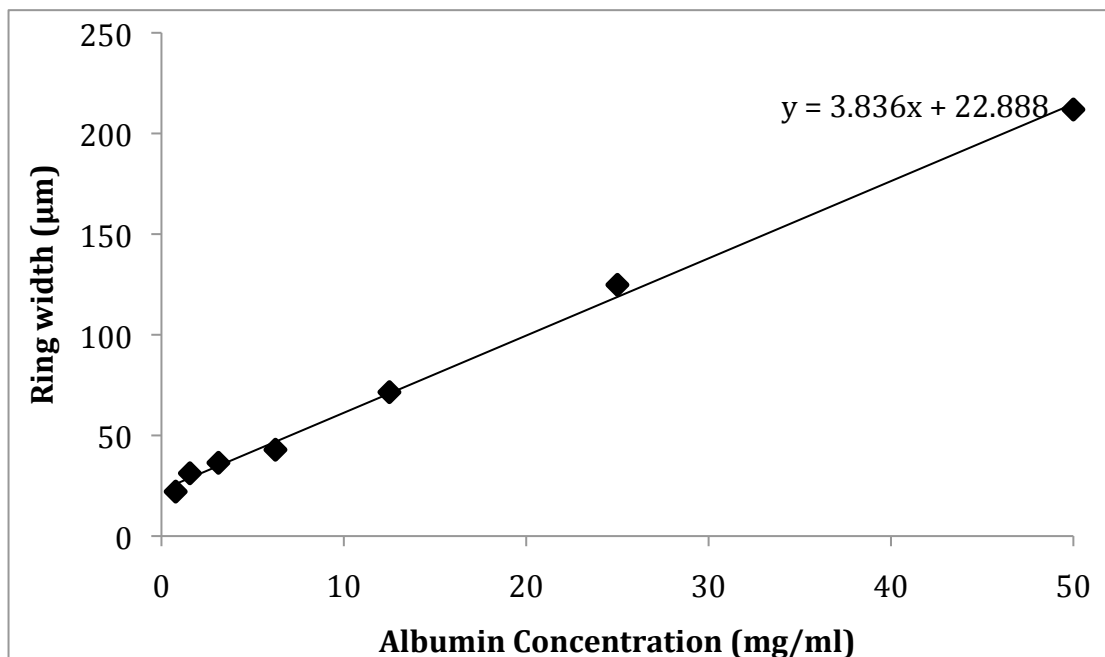


Figure 4.11. The Effect of Ring Width as Albumin Concentration Increases Across All Three Techniques

4.2.2 Amide I Peak Analysis

The amide I peak position was identified and recorded across the width of each ring as described in section 3.2.3. The mean and standard deviation were calculated and plotted for each droplet (Figure 4.12). At lower concentrations where cracks were visible in the rings, the spectra were removed from the line map data. The position of the amide I peak in the spectra measured in IR transmittance remains at 1656-1657 cm^{-1} with a standard deviation (SD) ranging from 0.88-1.33. As the FTIR spectrometer measures at intervals of two wavenumbers, 1SD is the equivalent of a variation around one data point. The position of the amide I peak is consistent with literature for an α -helical protein such as albumin (Chiriboga *et al*, 1998b & Sahu & Mordechai, 2005). In the Raman spectra, the mean amide I peak position is slightly more variable, with a range of 1656-1658 cm^{-1} in the three rings with the highest concentration with an SD ranging from 2.14-2.97. Again this is consistent with literature (Mahadevan-Jansen & Richards-Kortum, 1996). In rings A1, A2, A3 and A4 the mean amide I position and the SD shift considerably from 4.8-11.7 owing to there only being on average two measureable data points within these very thin rings, suggesting a limitation of this technique at concentrations below 12.5 mg/ml. The rings measured in the reference method, IR reflectance, show significant shifts to the amide I peak throughout all seven rings. The most concentrated ring A7 shows a significantly larger shift. Most of the data was removed from this ring owing to the large amount of cracks and thus only two data points were actually used, both of which showed the amide I peak to be at 1684 cm^{-1} .

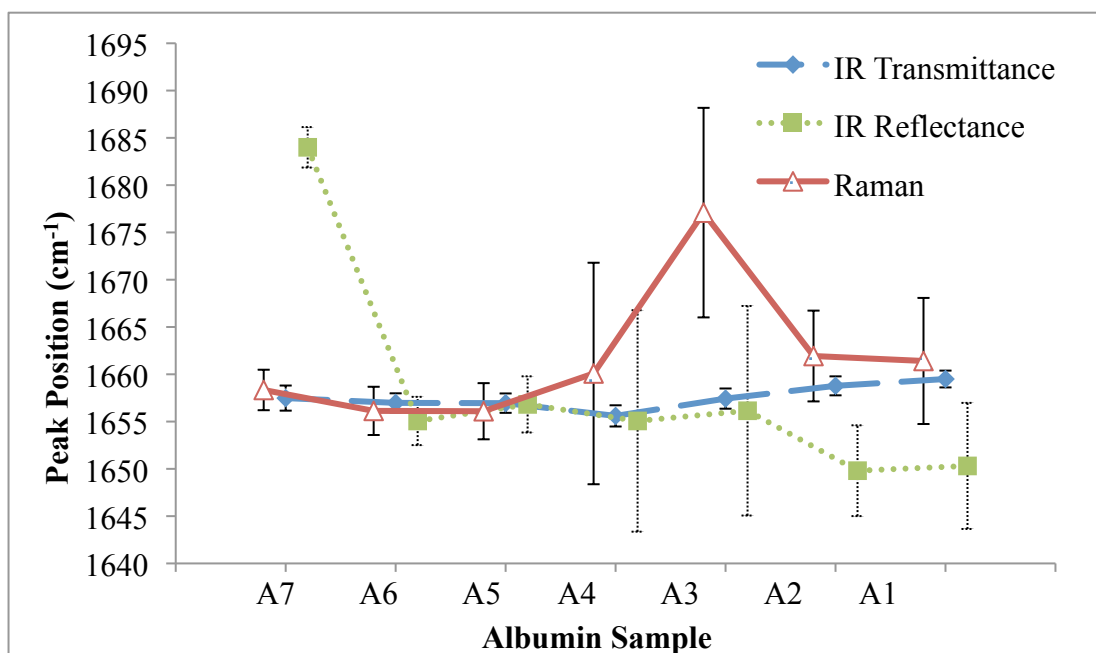


Figure 4.12. Average Amide I Peak Position Across All Rings Measured in Raman, IR Transmittance & IR Reflectance

These results indicate the change to the mean amide I peak position across each ring which does not reflect the peak position at different points across the ring. As the ring thickness is lower at the edges due to the curvature of the droplets, there may be effects from scattering that have not been identified. Line profile plots were created for each of the spectral maps obtained to show the amide I peak position and absorption/ intensity at each data point. The example shown in Figure 4.13 (for ring A7 in IR transmittance) shows the position of the amide I peak across the diameter of the ring. Looking at the highlighted region within the width of the ring, the amide I peak position remains at 1656-1658 cm^{-1} most of the way across the width of the ring, but starts to shift at the very edges where the concentration of albumin decreases, suggesting a small effect of scattering. Likewise, the absorption plot shows a sharp decrease at the edges of the ring suggesting a central region within the ring for optimum spectral acquisition. A similar picture was seen in all ring deposits in Raman and IR transmittance.

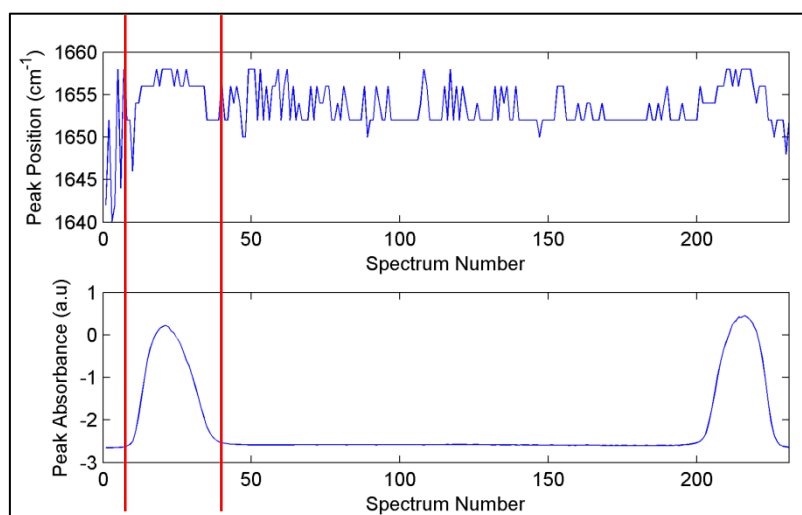


Figure 4.13. Line Profile Plot of Ring A7; showing the peak absorption and position of the amide I peak measured in IR transmittance

From these results it was necessary to identify where within the ring is appropriate to take spectra to eliminate spectral shifts and therefore to identify the minimum ring width for accurate results.

4.2.3. Determining Minimum Ring Width

The position of the amide I peak from each data point across the ring was used to calculate the SD. From both edges of the width of the ring one data point was removed at a time and the mean and standard deviation recalculated. This was done the entire way through the ring until the central data points were reached. These results were then plotted in order to determine the most precise and accurate area of the rings where spectra could be taken where the peak position was consistent and SD was low. The purpose of this was to deduce the minimum ring width required for reproducible results as well as to determine the distance from the edge for spectral acquisition that must be avoided in order to minimise the effects of scattering. Using data from ring A7 in both IR transmittance and Raman (Figure 4.14) the SD was plotted.

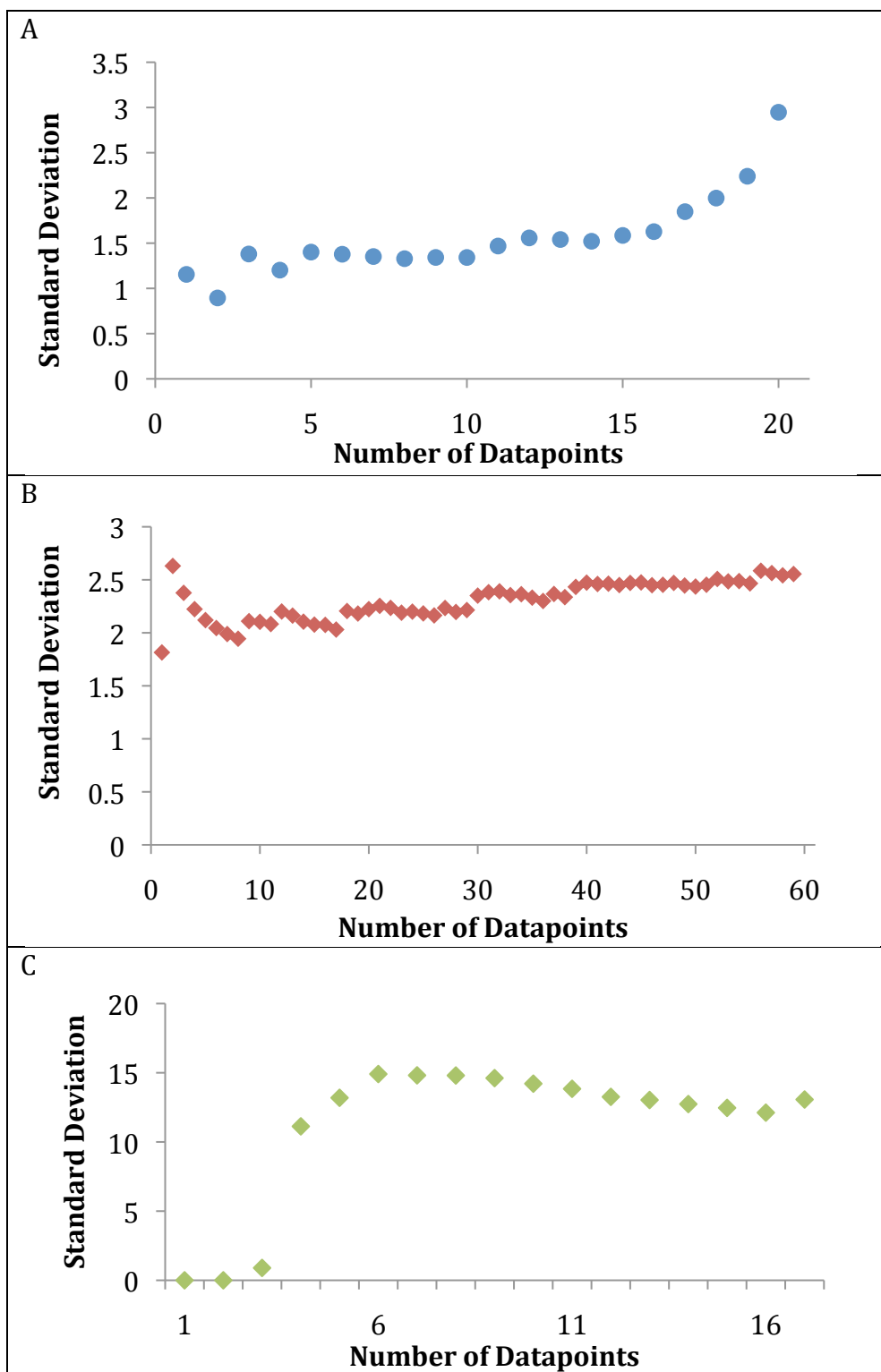


Figure 4.14. Measurement of Ring Width by Calculating Mean and Standard Deviation of the Amide I Peak at Data Points Across the Ring Width; A) IR transmittance, B) Raman, C) IR reflectance

In IR transmittance (Figure 4.14A) the SD is slightly raised at the edges of the ring. As two data points were removed from the edges the SD decreased to an acceptable level, below two. The width of the ring was 211.9 μm , which was divided by the number of pixels in the ring (81.5) to identify the number of microns per pixel (2.6 $\mu\text{m}/\text{pixel}$). This was then multiplied by the number of data points to identify how far from the edge of the ring was acceptable to obtain spectral measurements. Thus two data points equates to a distance of 5.2 μm . In the centre of the ring the SD remained low suggesting the results remain consistent regardless of the number of data points averaged.

The SD was on average higher in the Raman data (Figure 4.14B) but as the outer four data points were removed the SD dropped below 2.5 and remained between 2-2.5 throughout the ring. Using the same calculation, it was determined that spectral measurements should be taken at a minimum of 15.3 μm from the edge of the ring in order to ensure that any effects of scattering are avoided. In the centre of the ring the SD is increased suggesting that several measurements are required in order to obtain an accurate result. In IR reflectance (Figure 4.14C) the SD is very high due to the scattering.

4.2.4. Summary

Both Raman and IR transmittance spectroscopy were able to detect the albumin spectra at each concentration with little background interference. As the concentration of the rings decreased the absorbance of the IR spectra was slightly reduced but the signal to noise ratio remained low. Raman, however, was better at detecting the spectra of the more concentrated rings, with increased intensity and sensitivity. This technique is therefore limited to detecting protein concentrations above 25 mg/ml. However, the IR transmittance spectra remained consistent at each concentration, even below 25 mg/ml when the ring had deteriorated. It is important to note that in section 4.1 saturation occurred at 50 mg/ml and thus this would appear to be the maximum concentration of protein to use for IR transmittance. These results show the potential of DCDS as a tool for micro-analysis of biological samples with low protein concentration.

DCDS has been shown to be a useful tool for detecting low concentrations of proteins. Both FTIR and Raman results indicate the use of the technique to detect levels of albumin much lower than found in the blood. However, it has been shown that a minimum ring width is required for reducing the effects of scattering and that this ring width is affected by sample concentration. As the albumin concentration is decreased the ring width reduces thus lowering the number of possible spectral measurements. It also causes the ring structure to become less stable and crack.

4.3. Discussion

The method of DCDS has been utilised for protein analysis by several groups and enabled the detection of changes to protein conformation, binding and purity (Ortiz *et al*, 2006) as well as aided in the diagnosis of ocular disease (Filik & Stone, 2007). Tear fluid has been used to diagnose ocular abnormalities, not just from the spectra obtained, but from the fern-like drying patterns that arise upon deposition (Pearce & Tomlinson, 2000). In chapter 4 some of these ferning patterns were seen in the solutions containing fibrinogen, which could simply be due to the protein not dissolving in the solutions or from the interference of saline. However, if it is due to a characteristic of the protein itself, it could be used to identify the presence of an abnormality prior to spectroscopic analysis. Should these ferning patterns be seen in blood this could potentially be used in the diagnosis of leukaemia if it could be proven that these drying patterns related to the presence of disease.

The results obtained indicate that Raman is the better spectroscopic technique for higher protein concentrations. The higher the concentration, the thicker the ring therefore a better signal is achieved. However this may only be true to a certain degree as it has been identified that multiple pinning of the contact line can occur in higher concentration solutions (Esmonde-White *et al*, 2009). This may cause interference in the spectra, again limiting the DCDS method. In IR transmittance too thick a sample prevents light passing through thus reducing the quality of the spectra as they become saturated. IR transmittance should be

the method of choice for lower protein concentrations as indicated by the results. Although the absorbance at lower concentrations is reduced, the technique is sensitive enough to detect the peaks of interest.

In the next chapter the DCDS method is explored with blood samples. Although it has potential to be used with biological samples where unknown and uncontrollable concentrations of proteins are found, variability in ring formation is still a factor that is difficult to overcome. From this work and from the work of other research groups into the DCDS technique, forming droplets in which qualitative spectra can be obtained is fairly difficult and the best way to overcome this would be to use a series of dilutions and visually inspecting the droplets under a microscope prior to spectral acquisition. Rings that are too dilute and do not form properly should be avoided, as should droplets of too high concentration which suffer interference due to multiple pinning of the contact line. It has therefore been determined from this study that the following should be taken into consideration for optimising the DCDS method for leukaemia blood analysis:

- Sample preparation: make a series of dilutions of the sample in water and pipette these all onto a CaF₂ slide.
- Physical properties: prior to spectroscopic analysis ensure successful deposition has occurred. Check for bubbles and cracks. Minimum ring width must account for the distance at the edges of the ring where scattering may occur.
- Measurement: take IR spectral measurements at least 5.2 μm from the edges of the ring and from cracks in the ring. Take Raman spectral measurements at least 15.3 μm from the edges of the ring and from cracks in the ring.

5. RESULTS AND DISCUSSION: LEUKAEMIA ANALYSIS

This chapter outlines the results obtained from Raman and Fourier Transform Infrared (FTIR) Spectroscopy of blood specimens for the diagnosis of Leukaemia. For each sample, whole blood and plasma were deposited onto CaF₂ slides by drop coating deposition spectroscopy (DCDS) and whole blood and buffy coat by cytopspin centrifugation. The spectra were pre-processed, as described in section 3.4.4.4, to remove any differences in the data that may be mistaken for biochemical variances. Any saturated infrared (IR) spectra were also eliminated. The data was then analysed by univariate (mean spectra and difference spectra) and multivariate (principal component analysis (PCA) followed by linear discriminant analysis (LDA)) analysis to create a classification model. The performance of the classification model was assessed by leave one sample out cross validation (LOSOCV).

5.1 Screening Study

A total of 37 samples were obtained from 37 leukaemia patients and healthy volunteers. Eight different pathology groups were analysed; Chronic Lymphoblastic Leukaemia (Previously Untreated Patients) (CLL (PUP)), Chronic Myeloid Leukaemia (CML), Chronic Lymphoblastic Leukaemia (Previously Treated Patients) (CLL (PTP)), Plasma Cell Leukaemia (PCL), Acute Myeloid Leukaemia (AML), Mast Cell Leukaemia (MCL), Marginal Zone Lymphoma (MZL) and healthy control. The samples have been grouped into their pathologies in Table 5.1. Samples and spectra were obtained using the methods described in section 3.1 and initial screening studies were carried out to determine the best sampling methods.

Table 5.1. Samples Obtained

Pathology	Pathology label	Number of patients/ samples	Number of spectra
Chronic Lymphoblastic Leukaemia (Previously Untreated Patients)	CLL (PUP)	11	330
Chronic Myeloid Leukaemia	CML	2	60
Chronic Lymphoblastic Leukaemia (Previously Treated Patients)	CLL (PTP)	8	240
Plasma Cell Leukaemia	PCL	1	30
Acute Myeloid Leukaemia	AML	3	90
Mast Cell Leukaemia	MCL	1	30
Marginal Zone Lymphoma	MZL	2	60
Healthy	Healthy	9	270

Number of spectra is the number obtained for each sample type.

5.1.1. Morphological Analysis

Using the information obtained from the blood protein droplet studies in chapter 4, all samples were initially morphologically examined using the white light cameras in order to identify the most appropriate droplets for the two spectroscopic techniques. Poor quality samples that were small or had weak structures filled with cracks were also avoided.

5.1.2. Empirical Analysis

Using the parameters described in section 3.4.4.3, thirty Raman and thirty IR point spectra were collected from each of the four sample types from each patient. This number of spectra was selected based on power calculations from Babrahs work (2009). Measurements were taken from the central regions of the ring in the blood droplet samples, at least 5.2 μm (for IR) and 15.3 μm (for Raman) from the edge or any cracks. All FTIR and Raman spectra were loaded in

to Matlab and were normalised and mean centred. Some of the samples were too thick for IR transmittance analysis and thus the transmission of light was saturated causing the maximal absorbance value to be reached. This data was therefore subject to empirical screening to remove any of these saturated spectra from the dataset. This was done using the ratio of the amide I and amide II peaks, identified by the largest peak in the 1600-1700 cm^{-1} and 1500-1600 cm^{-1} regions respectively. The ratios were selected based on a mixture of references (Table 5.2).

Table 5.2. Amide I/ Amide II Ratios

Tissue or Substance	Amide I/ Amide II Ratio	Source
Globulin	1.27	Appendix E
Albumin	1.20	Appendix E
Fibrinogen	1.63	Appendix E
Non-Cancerous Lymph Node	1.38	Isabelle <i>et al</i> , 2008
Cancerous Lymph Node	1.74	Isabelle <i>et al</i> , 2008
Lymphoma Cell Lines	1.56	Babrah, 2009
Myeloid Cell Lines	1.57	Babrah, 2009
Lymphoid Cell Lines	1.70	Babrah, 2009
Karpus Cell Lines	1.49	Babrah, 2009

Histograms of the amide I/ amide II ratio of all the spectra for each of the four different pathology groups were then plotted (Figure 5.1). From the sources in Table 5.2 it is clear that the amide I/ amide II ratio is highly variable, depending on what is being measured. It can also be seen in the histograms that there is a difference between sample types, even between the blood cytopsin and the blood droplet. In order to eliminate saturated spectra, all of the data with a ratio of 1 were initially removed and the spectra reassessed for saturation. A ratio of 1 was chosen at first as all of the references indicated that the ratio of the two peaks was always above this value. If saturated spectra were still present in the data, all spectra with a ratio of 1.1 were removed and the spectra reassessed.

This was repeated until all saturated spectra were removed. For both buffy coat and blood cytopsin, all spectra with an amide I/ amide II ratio lower than 1.3 was removed. For the blood droplet and plasma, all spectra with an amide I/ amide II ratio lower than 1.1 was removed. The plasma IR spectra before and after saturated spectra removal are displayed in Figure 5.2.

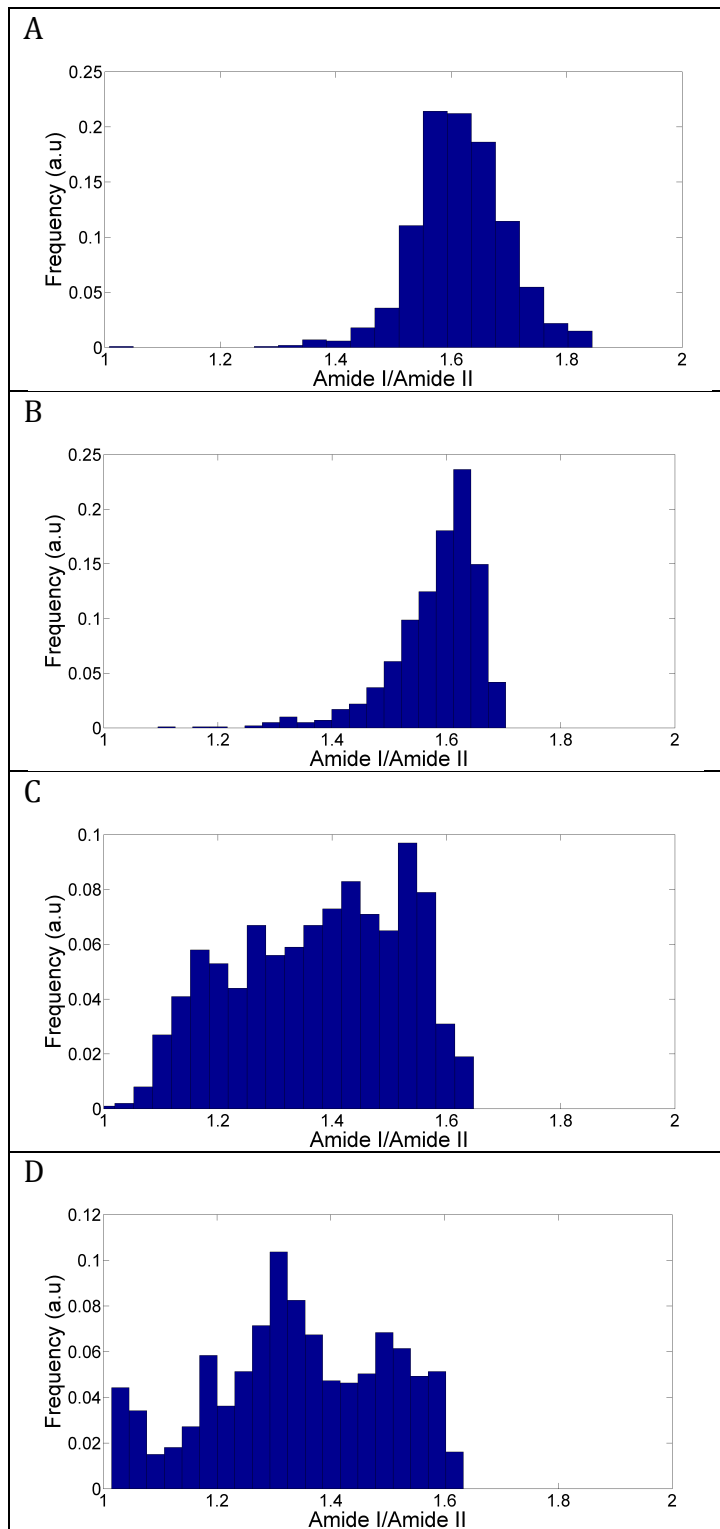


Figure 5.1. Histogram Plots of Amide I/ Amide II Ratios; A) Buffy Coat, B) Blood Cytospin, C) Blood Droplet, D) Plasma

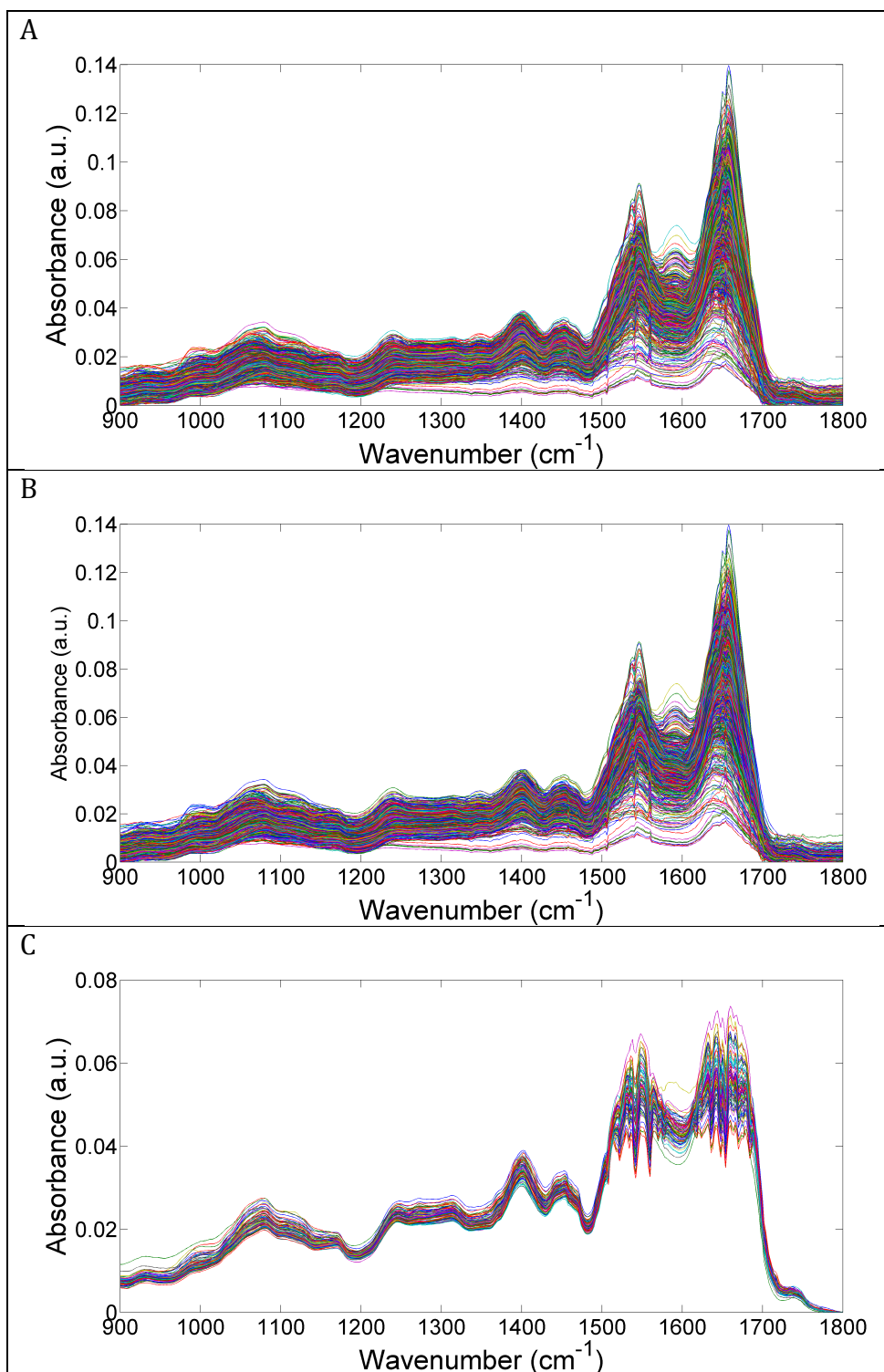


Figure 5.2. Plasma IR Spectra; A) before saturated spectra removed, B) after removal of saturated spectra, C) identified saturated spectra removed

Table 5.3. Number of Spectra Removed From Each Sample Due to Saturation

Sample	Pathology	Buffy Coat	Blood Cytospin	Blood Droplet	Plasma
A01	Healthy	0	0	0	0
A02	Healthy	0	0	0	0
A03	Healthy	0	0	0	0
A04	Healthy	0	0	0	0
A05	Healthy	0	3	0	0
A06	Healthy	0	0	0	0
A07	Healthy	0	0	0	0
A08	Healthy	0	0	0	0
A09	Healthy	0	0	0	0
1	CLL (PUP)	0	0	0	0
2	CML	0	0	0	10
3	CLL (PTP)	0	0	1	10
4	CLL (PUP)	0	0	0	9
5	CLL (PUP)	0	0	0	1
6	CLL (PUP)	0	2	11	8
7	CLL (PTP)	0	1	1	2
8	PCL	0	0	0	25
9	CLL (PUP)	0	0	0	23
10	AML	1	0	0	0
11	CLL (PTP)	0	0	0	0
12	MCL	0	0	1	0
13	CLL (PTP)	0	0	0	0
14	CML	0	0	0	0
15	CLL (PUP)	0	0	1	0
16	CLL (PUP)	0	0	3	0
17	MZL	0	0	0	0
18	AML	0	0	0	0
19	CLL (PTP)	0	0	0	3
20	AML	1	0	2	0
21	CLL (PUP)	0	0	3	0
22	CLL (PUP)	0	0	2	2
23	CLL (PUP)	0	0	0	0
24	CLL (PTP)	0	0	0	0
25	CLL (PUP)	0	0	0	0
26	MZL	0	2	0	0
27	CLL (PTP)	0	0	0	0
28	CLL (PTP)	0	0	0	0
TOTAL		2	8	25	103

A high proportion of the spectra from six plasma samples (highlighted in red in Table 5.3) were saturated and thus the entire data from those samples was eliminated. Several spectra from the remaining samples had a large peak at 1585 cm^{-1} between the amide I and amide II peaks. This has previously been identified as C=C of fibrinogen in blood plasma by Poon *et al* (2012) however, due to it not being present in all spectra in all samples it may be mistaken for pathological variation and thus the plasma IR data was excluded from all further analysis.

At this stage it was also decided that owing to the varying sizes of each pathology group that some were too small for multivariate statistical analysis. Therefore only CLL (PUP), CLL (PTP) and healthy groups were analysed further.

5.2. Mean Spectra

Mean IR and Raman spectra of the four sample types (blood cytospin, blood droplet, buffy coat cytospin and plasma droplets) were plotted using the healthy data to identify any differences in peak positions or absorbance. Peaks were identified according to a visual inspection and manual selection using an in house Matlab tool, which were then assigned according to literature. A full list of peak assignments for IR and Raman are displayed in Appendix B.

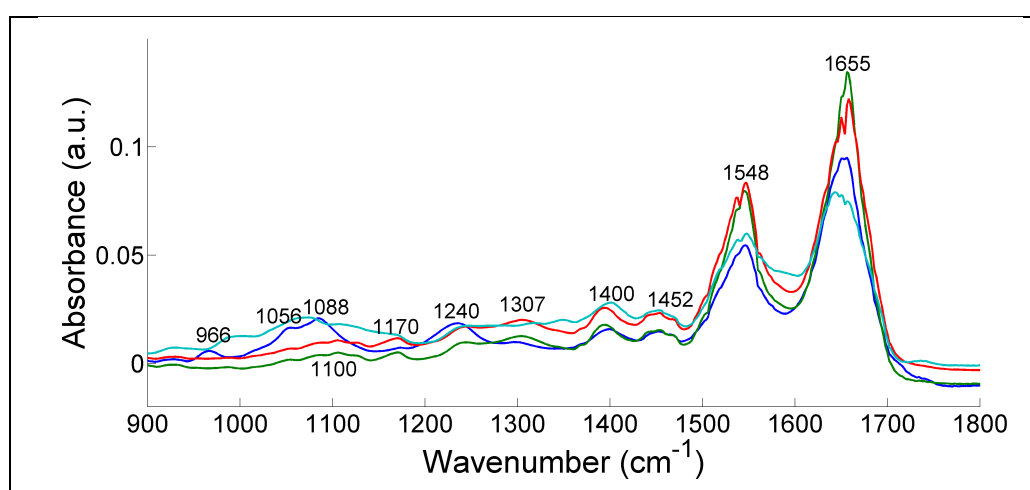


Figure 5.3. Mean Healthy IR Spectra of All Four Sample Types; dark blue: buffy coat, green: blood cytospin, red: blood droplet, light blue: plasma.

Across all four samples, peaks of varying intensities are observed at 1400 cm^{-1} , 1452 cm^{-1} , 1548 cm^{-1} and 1655 cm^{-1} corresponding to amino acids, lipids, amide II and amide I respectively (Lie *et al*, 1997, Liu *et al*, 2002 & Mourant *et al*, 2003b). In the buffy coat spectrum (dark blue), more prominent peaks can be seen at 966 cm^{-1} , 1056 cm^{-1} , 1088 cm^{-1} , and 1240 cm^{-1} corresponding to DNA, carbohydrates, nucleic acids and amide III respectively (Schultz *et al*, 1996, Lasch *et al*, 2002, Liu *et al*, 2002 & Erukhimovitch *et al*, 2006). A peak at 1307 cm^{-1} is also present in the buffy coat spectrum as well as in the blood droplet (red) and blood cytopsin (green) spectra. This peak corresponds to amide III (Chiriboga *et al*, 1998a). In addition, peaks at 1100 cm^{-1} and 1170 cm^{-1} , both relating to nucleic acids, are observed in the blood droplet and blood cytopsin spectra (Mourant *et al*, 2003b). The plasma data was also analysed to identify if any extra information may be lost by not using this sample group. The plasma spectra (light blue) did not appear to identify any other contributions from biochemical components apart from those already noted.

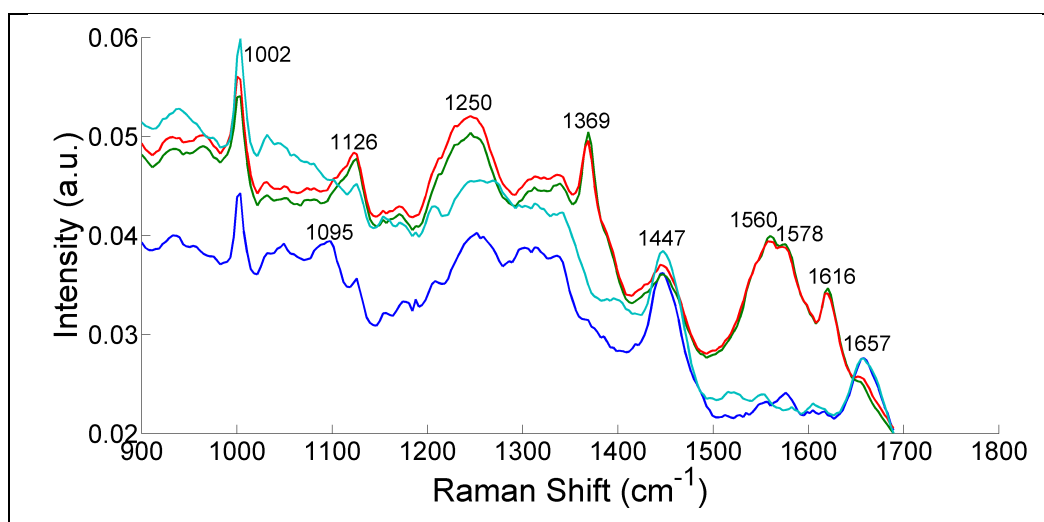


Figure 5.4. Mean Healthy Raman Spectra of All Four Sample Types; dark blue: buffy coat, green: blood cytopsin, red: blood droplet, light blue: plasma.

Similarly there are peaks observed in the mean Raman spectra of all four sample types. At 1002 cm^{-1} , 1250 cm^{-1} and 1447 cm^{-1} peaks are present which relate to phenylalanine (an amino acid), amide III and CH_2 vibrations from lipids/proteins respectively (Stone *et al*, 2002 & Pully *et al*, 2010). In the buffy coat spectra (dark blue) peaks can be seen at 1095 cm^{-1} relating to PO_2^- vibrations in DNA/ RNA and 1657 cm^{-1} relating to the C=O stretch of amide I (Lyng *et al*, 2007 & Poon *et al*, 2012). The amide I peak is also present in the plasma spectra (dark blue) but is masked in both the blood cytopsin (green) and blood droplet spectra (red) by a series of peaks at 1560 cm^{-1} , 1578 cm^{-1} (pyrimidine ring/ haem protein), and 1616 cm^{-1} (C=C vibrations of the amino acids tryptophan/tyrosine) (Stone *et al*, 2002 & Poon *et al*, 2012). The peaks at 1560 cm^{-1} and 1369 cm^{-1} cannot be identified from the literature. However peaks corresponding to amino acids have been identified in similar regions, such as tryptophan at 1361 cm^{-1} and 1560 cm^{-1} and tyrosine at 1375 cm^{-1} (Fredericks, 1995, Mahadevan-Jansen & Richards-Kortum, 1996 & Pully *et al*, 2010). Another peak at 1126 cm^{-1} can be seen in the blood cytopsin and blood droplet data only which is caused by C-C/ C-N vibrations of lipids/ proteins (Stone *et al*, 2002 & Bonnier *et al*, 2012).

Mean spectra of the three pathologies, healthy, CLL (PUP) and CLL (PTP) were also plotted. The mean buffy coat IR spectra of the three pathologies is shown in figure 5.5. The peaks of the main biochemical component have been highlighted. It can be seen that there are some subtle differences between the spectra, caused by variations in the absorption and positions of the peaks. However these differences are very difficult to identify without the use of further statistical analysis and thus the mean spectra for the other datasets are shown in Appendix F.

The peaks at 966 cm^{-1} and 1088 cm^{-1} are caused by the PO_2^- stretching vibrations of the nucleic acids in DNA (Schultz *et al*, 1996 & Liu *et al*, 2002). The peaks at 1056 cm^{-1} and 1452 cm^{-1} are attributed to vibrations from the C-OH of carbohydrates (Erukhimovitch *et al*, 2006) and the asymmetric deformation of CH_2 and CH_3 of lipids (Liu *et al*, 1997) respectively. The remaining peaks are caused by the absorption modes of proteins, with the most prominent peaks

observed in the spectra due to the C=O stretching vibrations of amide I at 1656 cm^{-1} and the N-H deformation of amide II at 1546 cm^{-1} (Liu *et al*, 2002). Other protein peaks are observed at 1240 cm^{-1} due to the C-N stretch and N-H bending of amide III (Lasch *et al*, 2002) and at 1400 cm^{-1} due to the bending vibrations of NH_2 in amino acids (Mourant *et al*, 2003b).

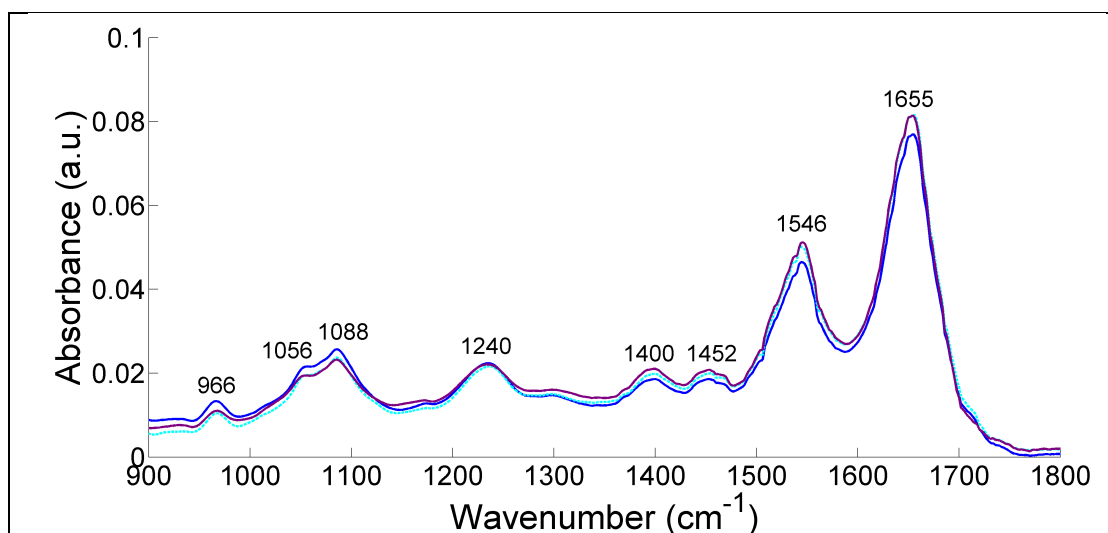


Figure 5.5. Mean Buffy Coat IR Spectra of the Three Pathologies; purple: healthy, dark blue: CLL (PUP), light blue: CLL (PTP)

The mean buffy coat Raman spectra of the three pathologies are shown in Figure 5.6. Again, the mean spectra for the other datasets are shown in Appendix F. There are several differences in the Raman spectra compared to IR. There are many more peaks and the intensities of the biochemical components differ due to the different sensitivities of the two techniques to certain functional groups. Raman has stronger signals from C=C and aromatic rings and IR has stronger signals from polar groups such as C=O, C-O, C-H and O-H. Due to the number of peaks seen the peak assignments are displayed in Table 5.4.

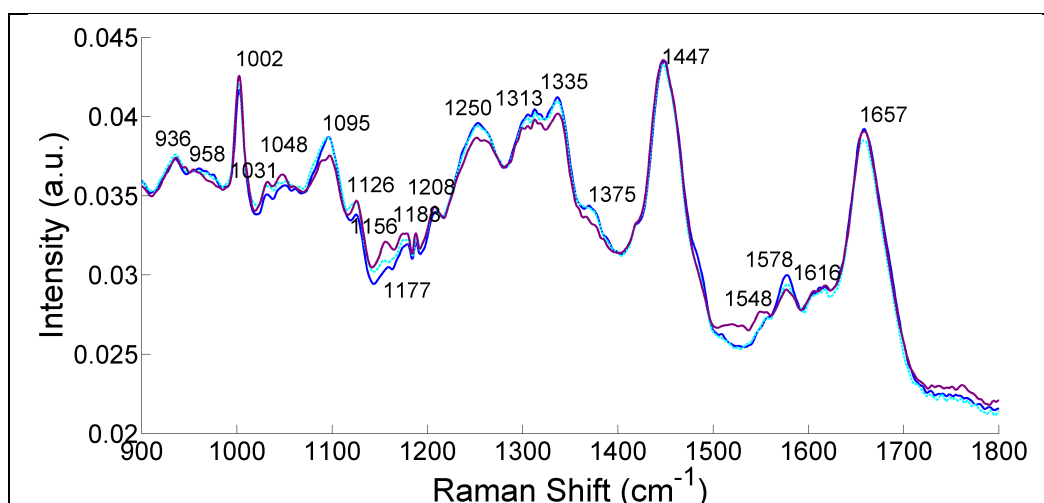


Figure 5.6. Mean Buffy Coat Raman Spectra of the Three Pathologies;
purple: healthy, dark blue: CLL (PUP), light blue: CLL (PTP)

Table 5.4. Raman Peak Assignments

Peak (cm ⁻¹)	Assignment	Reference
936	C-C vibrations α -helical protein backbone	Mahadevan-Jansen & Richards-Kortum 1996
958	CH ₃ deformation	Poon <i>et al</i> , 2012
1002	Symmetric ring breathing phenylalanine	Stone <i>et al</i> , 2002
1031	C-H vibrations phenylalanine	Poon <i>et al</i> 2012
1048	Weak C=O stretch Glycogen	Kendall, 2002
1095	PO ₂ ⁻ DNA/RNA	Mahadevan-Jansen & Richards-Kortum 1996
1126	C-N stretch protein	Bonnier <i>et al</i> , 2012
1156	C-N stretch proteins	Fredericks, 1995
1177	NH ₃	Bonnier <i>et al</i> , 2012
1188		
1208	C ₆ H ₅ tryptophan and phenylalanine	Poon <i>et al</i> , 2012
1250	C-N vibrations amide III	Pully <i>et al</i> , 2010
1313	CH ₂ CH ₃ deformation lipids	Stone <i>et al</i> , 2002
1335	CH ₂ CH ₃ deformation nucleic acids/ proteins	Stone <i>et al</i> , 2002 & Pully <i>et al</i> , 2010
1375	Ring breathing modes DNA	Pully <i>et al</i> , 2010
1447	CH ₂ bending proteins/ lipids	Pully <i>et al</i> , 2010
1548	N-H deformation/ C-N vibration amide II/ tryptophan	Lyng <i>et al</i> , 2007 & Stone <i>et al</i> , 2002
1578	Pyrimidine ring/ haem protein	Stone <i>et al</i> , 2002
1616	C=C vibrations tyrosine/ tryptophan	Poon <i>et al</i> , 2012
1657	C=O stretch amide I/ C=C stretch lipid	Poon <i>et al</i> , 2012

5.3. Difference Spectra

Difference spectra were generated to identify biochemical differences between the three pathologies. Using Matlab, the difference spectra were calculated by subtracting the mean spectra of one pathology from the mean spectra of another with differences represented by negative peaks below zero and positive peaks above zero respectively. An in house peak assignment tool was used to identify the largest biochemical contributions in the spectra and the peak tables in Appendix B were used to identify them. Figures 5.7 – 5.13 show the difference spectra for all the datasets. Due to the large number of peaks in these complex spectra, the largest differences between the pathologies have been described in Tables 5.5 – 5.11. Throughout, differences can be seen in the protein content as represented by peaks in the amide I, II and III bands. Changes in nucleic acid and DNA due to increased cell proliferation are also observed. Other differences seen are due to fatty acids and phospholipids, which make up the cell membrane, and carbohydrates, the cells energy source. Thus cellular changes due to cancer can be identified in the spectra. There does not appear to be much biochemical variation between the two cancerous pathologies, CLL (PUP) and CLL (PTP). There is a slightly higher DNA contributions in CLL (PUP) whereas amide I and amide II have higher contributions in CLL (PTP).

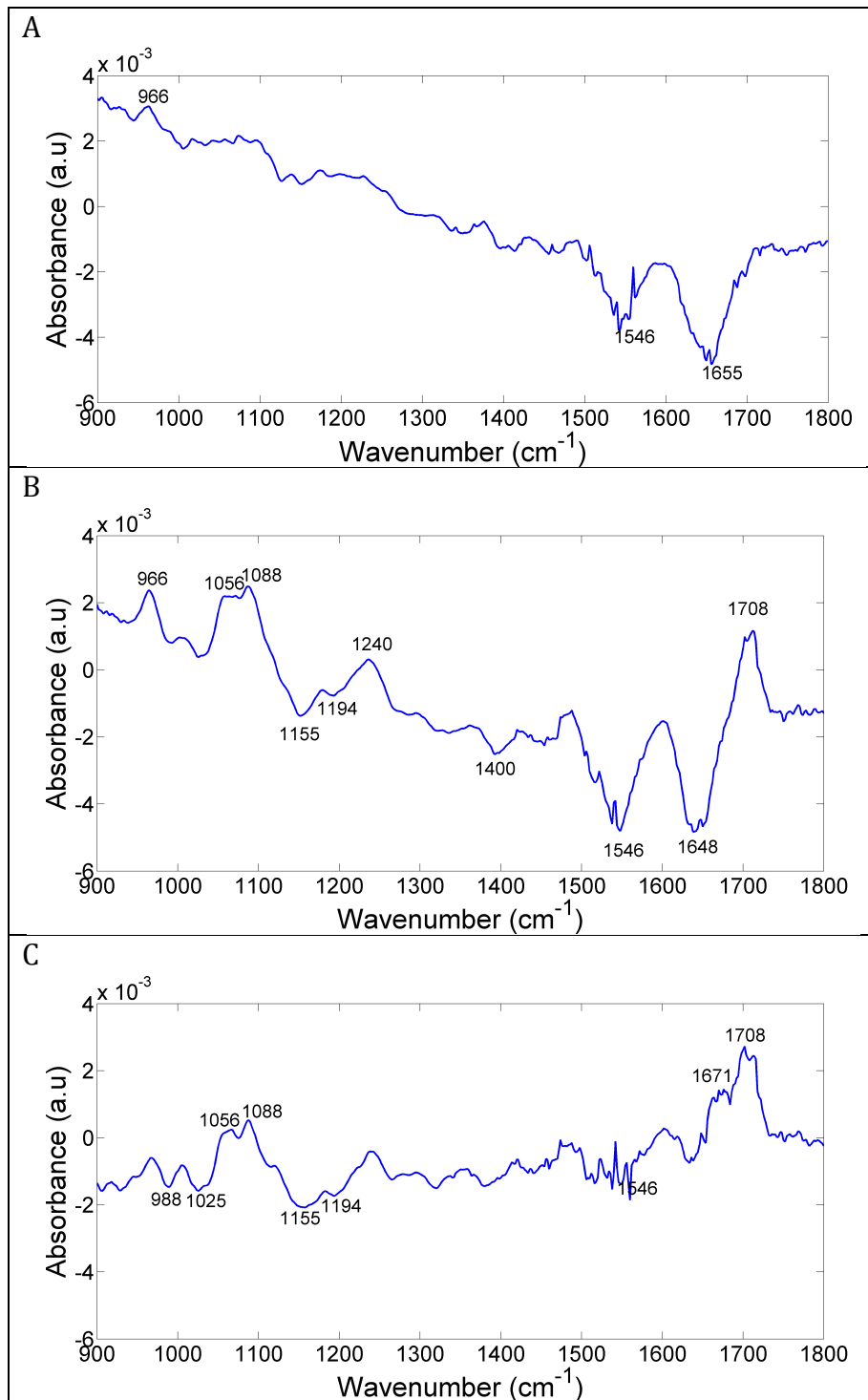


Figure 5.7. Difference Spectra for Buffy Coat IR; A) CLL (PUP) – CLL (PTP), B) CLL (PUP) – Healthy, C) CLL (PTP) - Healthy

Table 5.5. Assignment of Difference Spectra Peaks for Buffy Coat IR

	A	B	C
	CLL (PUP)	CLL (PUP)	CLL (PTP)
Positive peaks (cm⁻¹)	966 DNA	966 DNA 1056 C-OH carbohydrates 1088 PO ₂ ⁻ DNA (nucleic acids) 1240 C-N stretch/ N-H bending of amide III 1708 C-CO RNA	1056 C-OH carbohydrates 1088 PO ₂ ⁻ DNA (nucleic acids) 1671 C=O amide I 1708 C-CO RNA
	CLL (PTP)	Healthy	Healthy
Negative peaks (cm⁻¹)	1546 N-H amide II 1655 C=O amide I	1155 C-OH carbohydrates 1194 PO ₂ ⁻ asymmetric stretch 1400 NH ₂ in amino acids 1546 N-H amide II 1648 C=O amide I	988 RNA 1025 C-OH 1155 C-OH carbohydrates 1194 PO ₂ ⁻ asymmetric stretch 1546 N-H C=O amide II

A) CLL (PUP) – CLL (PTP), B) CLL (PUP) – Healthy, C) CLL (PTP) – Healthy

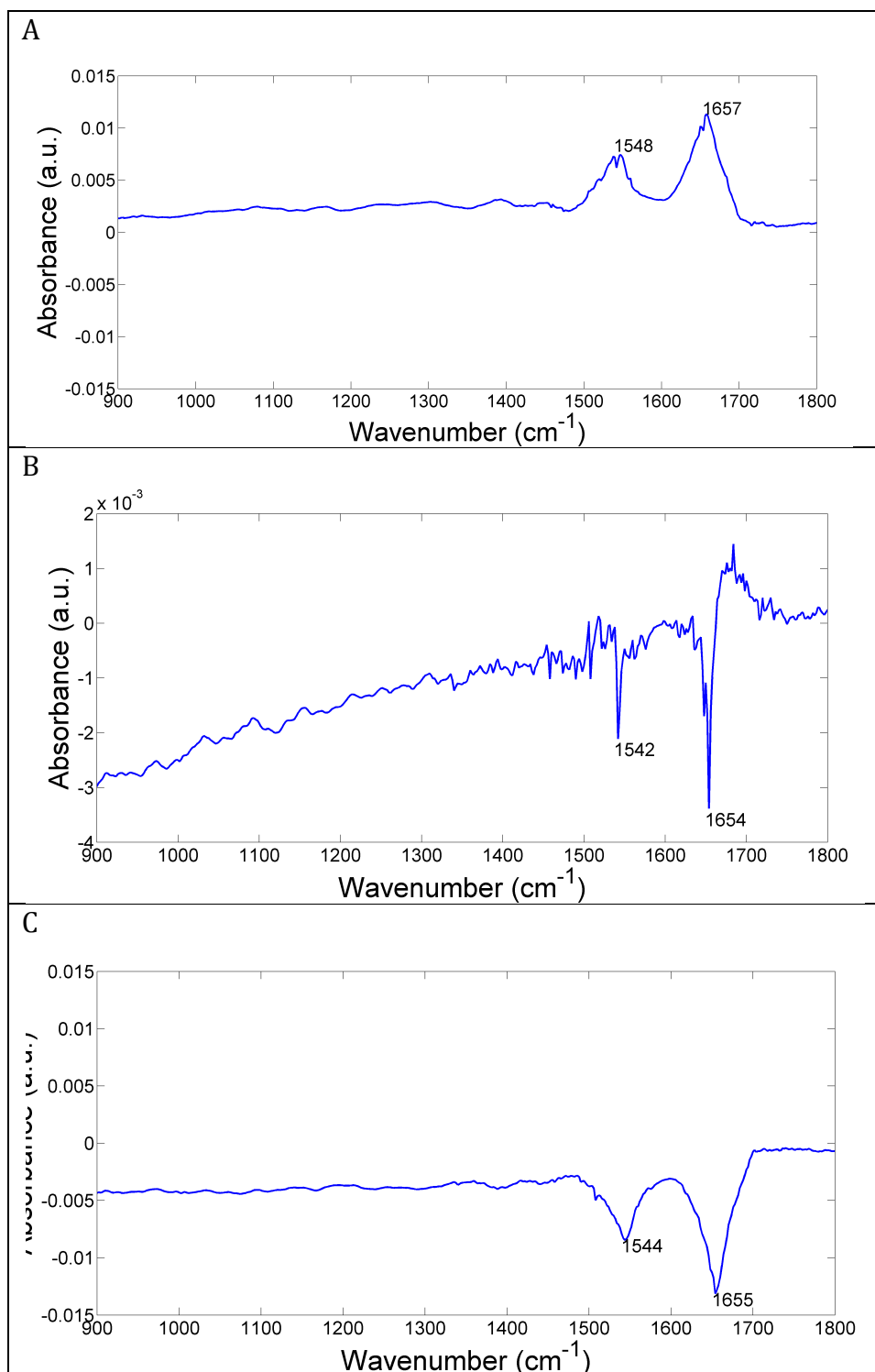


Figure 5.8. Difference Spectra for Blood Cytospin IR; A) CLL (PUP) - CLL (PTP), B) CLL (PUP) - Healthy, C) CLL (PTP) - Healthy

Table 5.6. Assignment of Difference Spectra Peaks for Blood Cytospin IR

	A	B	C
	CLL (PUP)	CLL (PUP)	CLL (PTP)
Positive peaks (cm⁻¹)	1548 N-H amide II 1657 C=O amide I		
	CLL (PTP)	Healthy	Healthy
Negative peaks (cm⁻¹)		1542 N-H amide II 1654 C=O amide I	1544 N-H amide II 1655 C=O amide I

A) CLL (PUP) – CLL (PTP), B) CLL (PUP) – Healthy, C) CLL (PTP) – Healthy

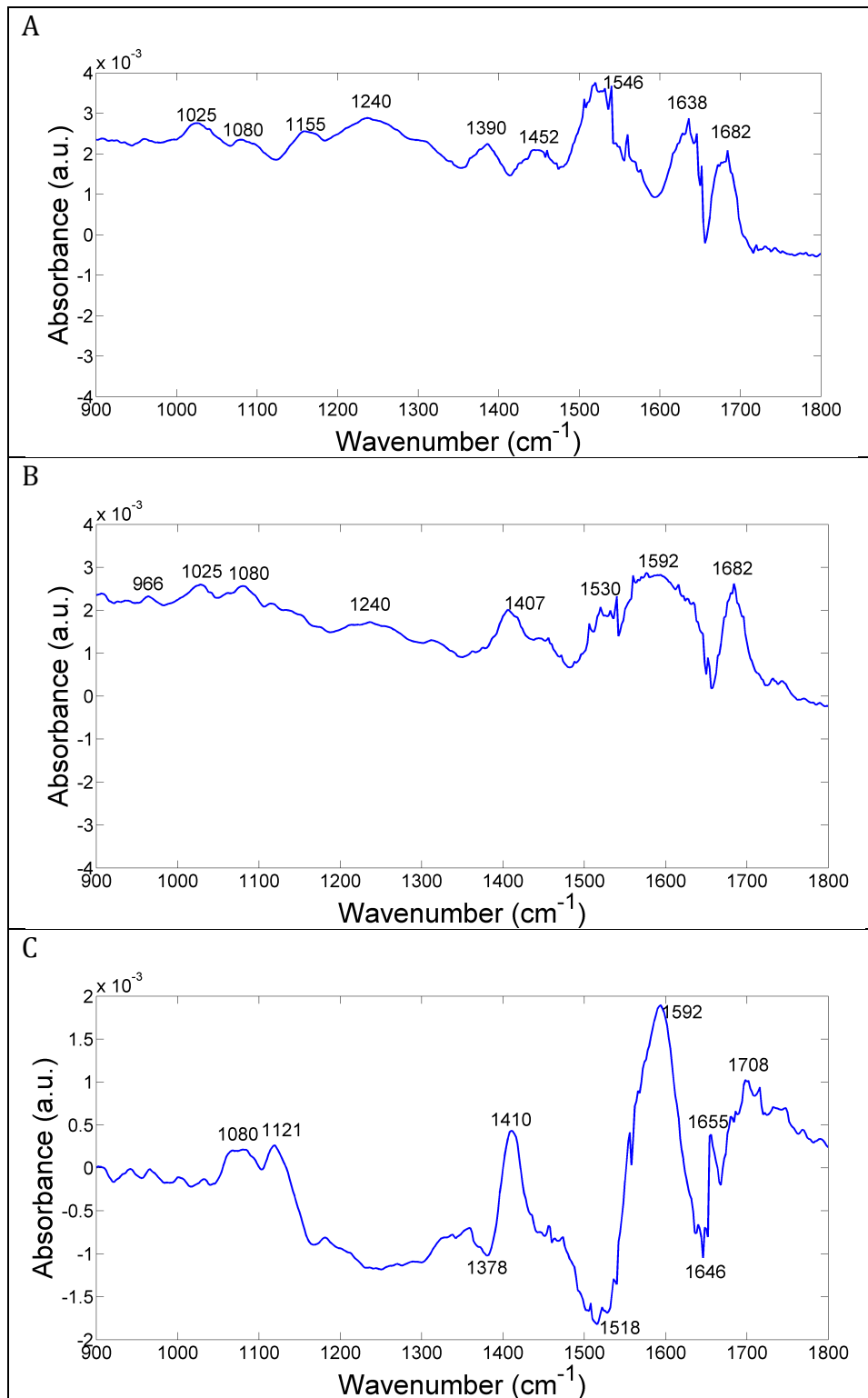


Figure 5.9. Difference Spectra for Blood Droplet IR; A) CLL (PUP) – CLL (PTP), B) CLL (PUP) – Healthy, C) CLL (PTP) - Healthy

Table 5.7. Assignment of Difference Spectra Peaks for Blood Droplet IR

	A	B	C
	CLL (PUP)	CLL (PUP)	CLL (PTP)
Positive peaks (cm⁻¹)	1025 C-OH carbohydrates 1080 PO ₂ ⁻ DNA (nucleic acids) 1155 C-OH carbohydrates 1240 C-N stretch/ N-H bending of amide III 1390 1452 CH ₃ CH ₂ fatty acids/ phospholipids 1546 N-H amide II 1638 1682 C=O amide I	966 DNA 1025 C-OH carbohydrates 1080 PO ₂ ⁻ DNA (nucleic acids) 1240 C-N stretch/ N-H bending of amide III 1407 CH ₃ CH ₂ fatty acids/ phospholipids 1530 N-H amide II 1592 1682 C=O amide I	1080 PO ₂ ⁻ DNA (nucleic acids) 1121 PO ₂ ⁻ DNA (nucleic acids) 1410 CH ₃ CH ₂ fatty acids/ phospholipids 1592 1655 C=O amide I 1708 C-CO RNA
	CLL (PTP)	Healthy	Healthy
Negative peaks (cm⁻¹)			1378 CH ₃ CH ₂ 1518 N-H tyrosine 1544 N-H amide II 1646 C=O amide I

A) CLL (PUP) – CLL (PTP), B) CLL (PUP) – Healthy, C) CLL (PTP) – Healthy

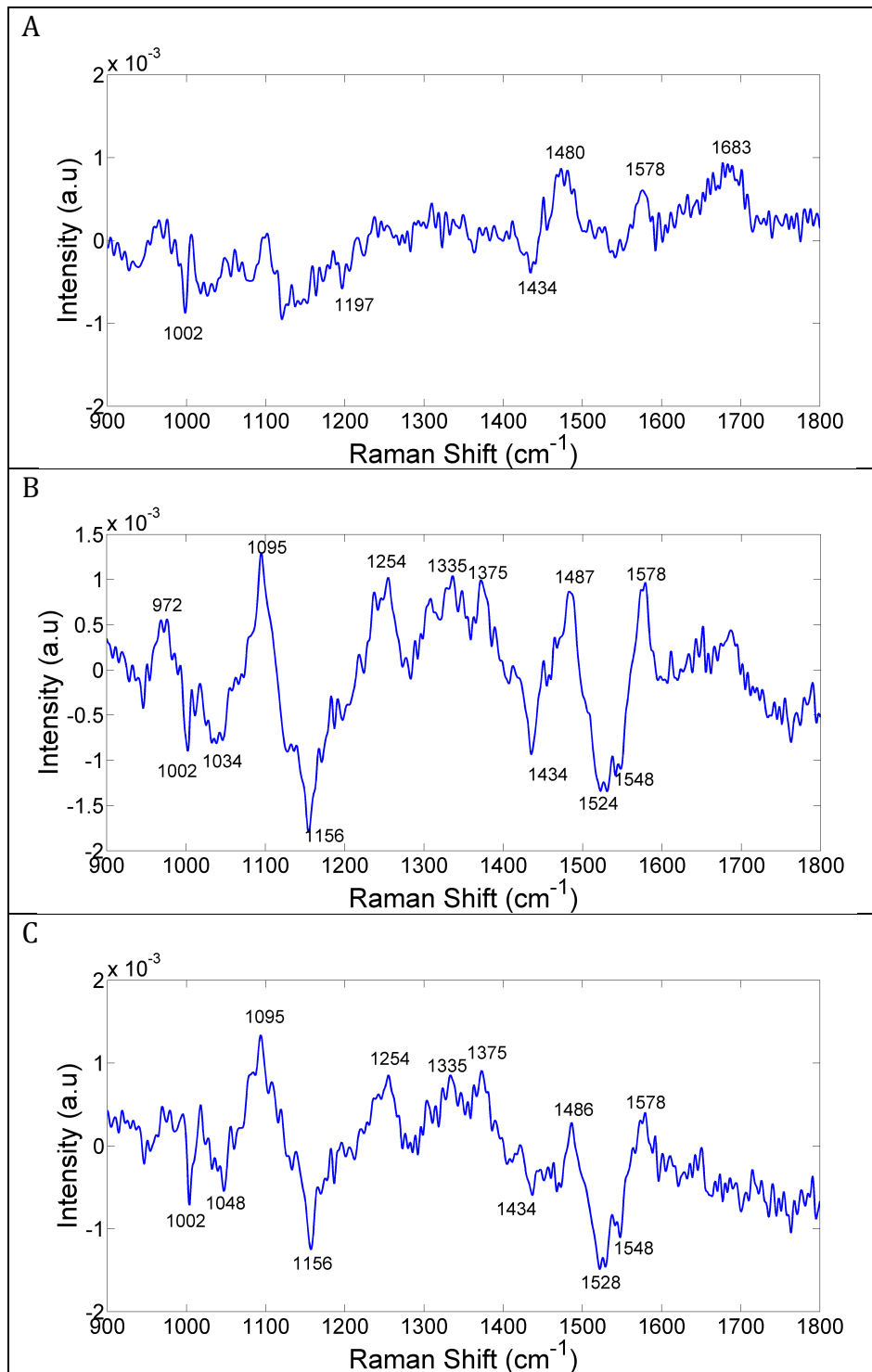


Figure 5.10. Difference Spectra for Buffy Coat Raman; A) CLL (PUP) – CLL (PTP), B) CLL (PUP) – Healthy, C) CLL (PTP) - Healthy

Table 5.8. Assignment of Difference Spectra Peaks for Buffy Coat Raman

	A	B	C
	CLL (PUP)	CLL (PUP)	CLL (PTP)
Positive peaks (cm⁻¹)	<p>1480 purine bases</p> <p>1578 N-H amide II</p> <p>1683 C=O amide I</p>	<p>972 C-C proline ring</p> <p>1095 PO₂⁻ DNA, RNA</p> <p>1254 C-N stretch/ N-H bending of amide III</p> <p>1335 CH₃CH₂ protein/ nucleic acid</p> <p>1375 ring breathing DNA</p> <p>1487 guanine DNA</p> <p>1578 N-H amide II</p>	<p>1095 PO₂⁻ DNA, RNA</p> <p>1254 C-N stretch/ N-H bending of amide III</p> <p>1335 CH₃CH₂ protein/ nucleic acid</p> <p>1375 ring breathing DNA</p> <p>1487 guanine DNA</p> <p>1578 N-H amide II</p>
	CLL (PTP)	Healthy	Healthy
Negative peaks (cm⁻¹)	<p>1002 phenylalanine</p> <p>1197 tyrosine</p> <p>1434 CH₂ lipids</p>	<p>1002 phenylalanine</p> <p>1034 C-C</p> <p>1156 C-N proteins</p> <p>1434 CH₂ lipids</p> <p>1524 C=C carotenoids</p> <p>1548 N-H amide II/ tryptophan</p>	<p>1002 phenylalanine</p> <p>1048 glycogen</p> <p>1156 C-N proteins</p> <p>1434 CH₂ lipids</p> <p>1528 C=C carotenoids</p> <p>1548 N-H amide II/ tryptophan</p>

A) CLL (PUP) – CLL (PTP), B) CLL (PUP) – Healthy, C) CLL (PTP) – Healthy

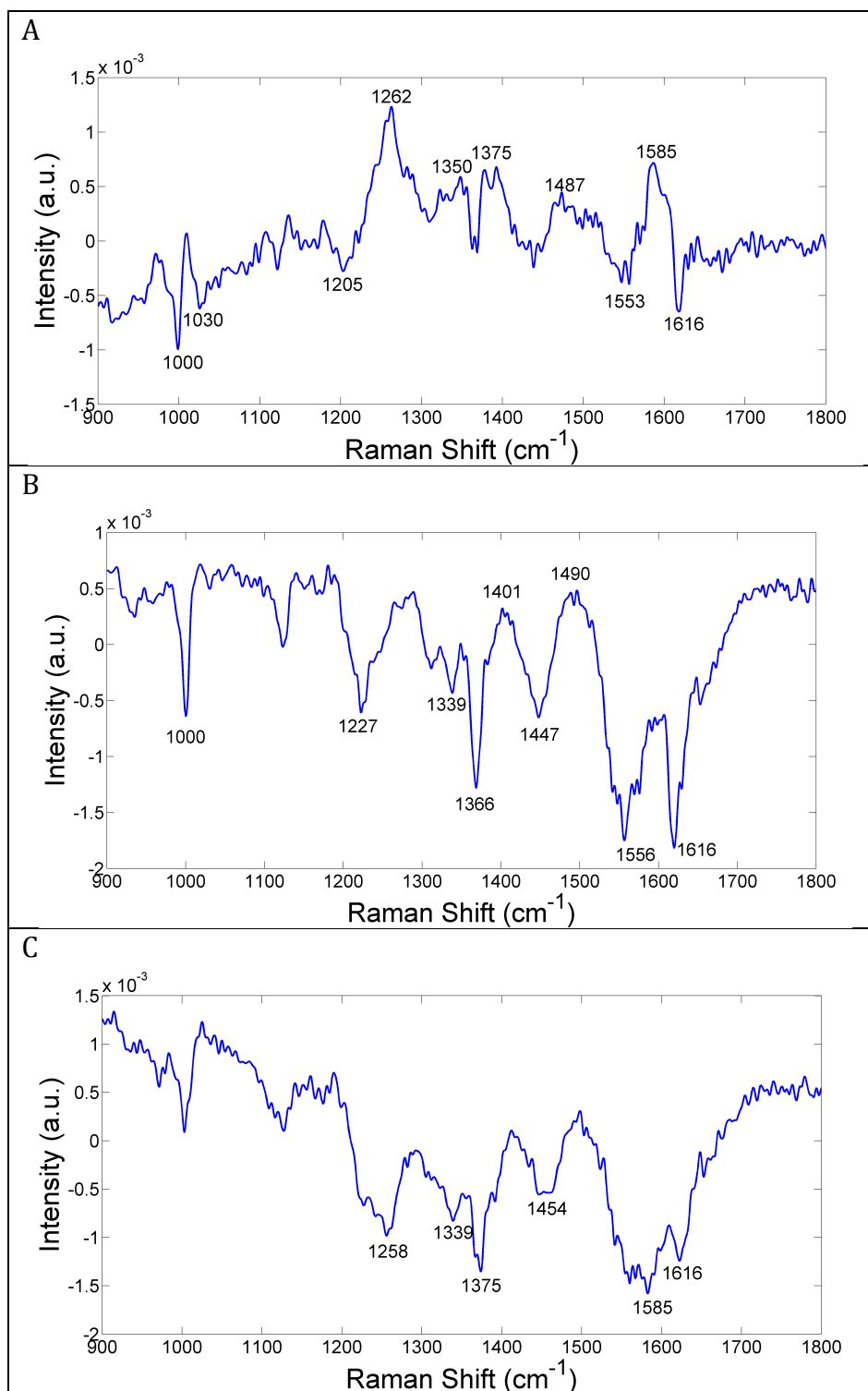


Figure 5.11. Difference Spectra for Blood Cytospin Raman; A) CLL (PUP) – CLL (PTP), B) CLL (PUP) – Healthy, C) CLL (PTP) - Healthy

Table 5.9. Assignment of Difference Spectra Peaks for Blood Cytospin Raman

	A	B	C
	CLL (PUP)	CLL (PUP)	CLL (PTP)
Positive peaks (cm⁻¹)	1262 C-N stretch/ N-H bending of amide III 1350 glucose 1375 ring breathing modes DNA 1487 guanine DNA 1585 C=C	1401 CH ₃ protein 1490 guanine DNA	
	CLL (PTP)	Healthy	Healthy
Negative peaks (cm⁻¹)	1000 glucose 1030 C-H phenylalanine 1205 tyrosine/phenylalanine 1553 tryptophan 1616 C=C tyrosine/tryptophan	1000 glucose 1227 C-N stretch/ N-H bending of amide III 1339 C-H 1366 CH ₂ 1447 CH ₂ protein/lipid 1556 tryptophan 1616 C=C tyrosine/tryptophan	1258 C-N stretch/ N-H bending of amide III 1339 C-H 1375 ring breathing modes DNA 1454 CH ₃ proteins 1585 C=C 1616 C=C tyrosine/tryptophan

A) CLL (PUP) – CLL (PTP), B) CLL (PUP) – Healthy, C) CLL (PTP) – Healthy

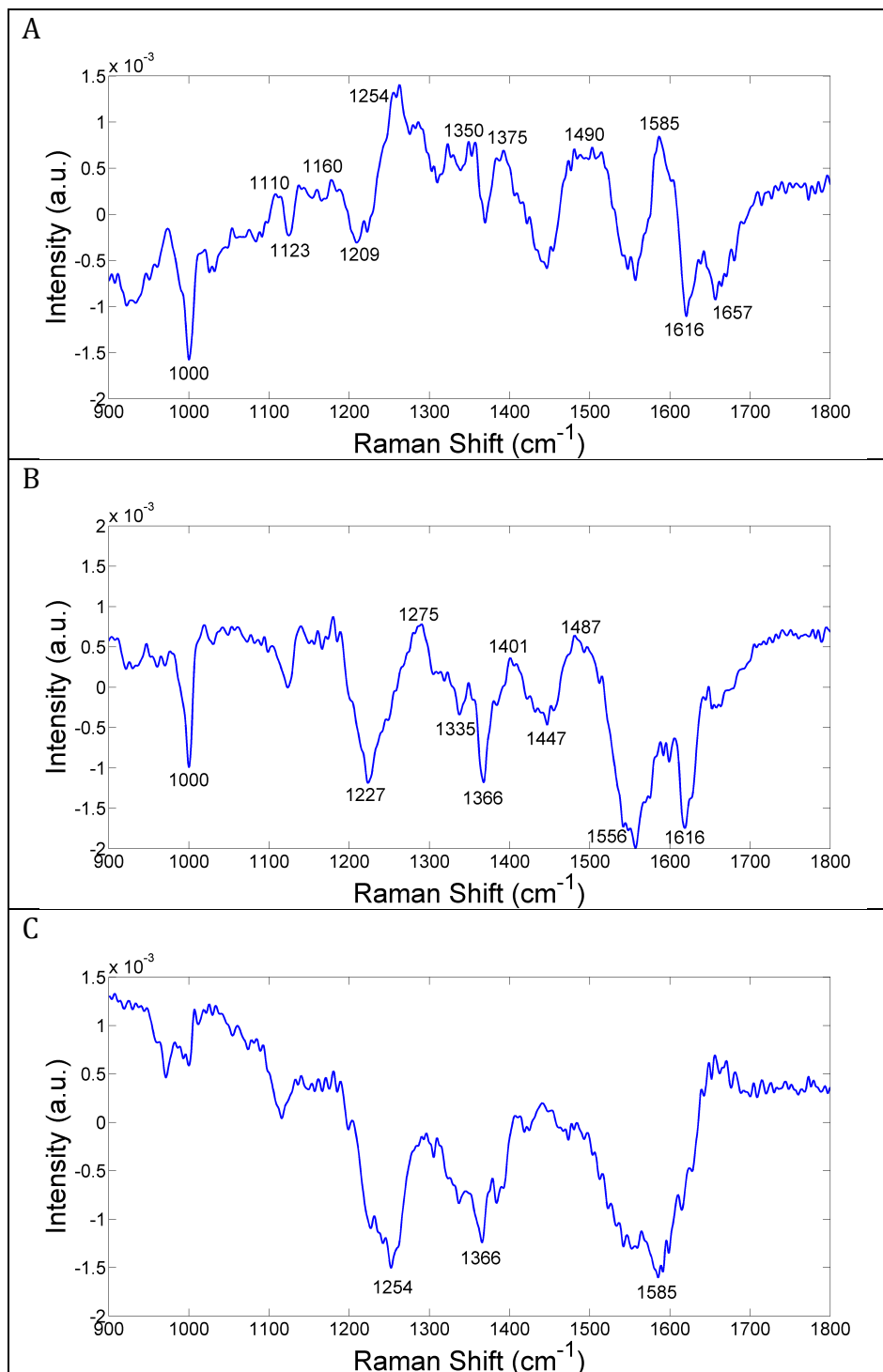


Figure 5.12. Difference Spectra for Blood Droplet Raman; A) CLL (PUP) – CLL (PTP), B) CLL (PUP) – Healthy, C) CLL (PTP) - Healthy

Table 5.10. Assignment of Difference Spectra Peaks for Blood Droplet Raman

	A	B	C
	CLL (PUP)	CLL (PUP)	CLL (PTP)
Positive peaks (cm⁻¹)	1110 C-C 1160 carotenoid 1254 C-N stretch/ N-H bending of amide III 1350 glucose 1375 ring breathing modes DNA 1490 guanine DNA 1585 C=C	1275 ring breathing modes DNA 1401 CH ₃ protein 1487 guanine DNA	
	CLL (PTP)	Healthy	Healthy
Negative peaks (cm⁻¹)	1000 glucose 1123 glucose 1209 C-C ₆ H ₅ tryptophan 1616 C=C tyrosine/tryptophan 1657 C=O amide I	1000 glucose 1227 C-N stretch/ N-H bending of amide III 1335 CH ₂ CH ₃ nucleic acids/ proteins 1366 CH ₂ 1447 CH ₂ protein/lipid 1556 tryptophan 1616 C=C tyrosine/tryptophan	1254 C-N stretch/ N-H bending of amide III 1366 CH ₂ CH ₃ nucleic acids/ proteins 1585 C=C

A) CLL (PUP) – CLL (PTP), B) CLL (PUP) – Healthy, C) CLL (PTP) – Healthy

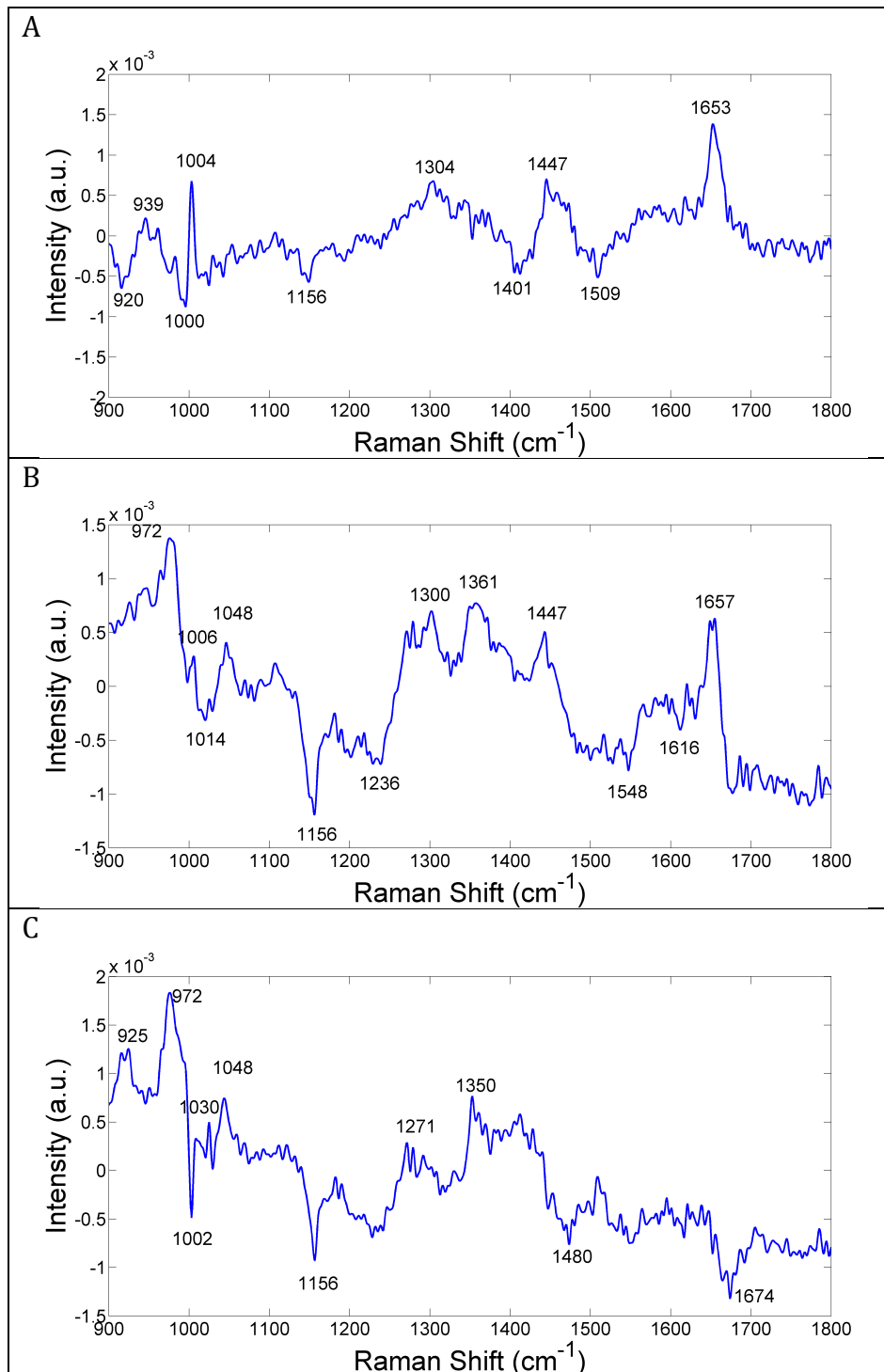


Figure 5.13. Difference Spectra for Plasma Raman; A) CLL (PUP) – CLL (PTP), B) CLL (PUP) – Healthy, C) CLL (PTP) - Healthy

Table 5.11. Assignment of Difference Spectra Peaks for Plasma Raman

	A	B	C
	CLL (PUP)	CLL (PUP)	CLL (PTP)
Positive peaks (cm⁻¹)	939 C-CN 1004 phenylalanine 1304 CH ₃ CH ₂ twisting 1447 CH ₂ protein/lipid 1653 C=O amide I	972 C-C proline 1006 phenylalanine 1048 C=O glycogen 1300 C-H lipid/ fatty acid 1361 tryptophan 1447 CH ₂ protein/lipid 1653 C=O amide I	925 proline ring 972 C-C proline 1030 phenylalanine 1048 C=O glycogen 1271 C-N stretch/ N-H bending of amide III 1350 glucose
	CLL (PTP)	Healthy	Healthy
Negative peaks (cm⁻¹)	920 glucose 1000 glucose 1156 C-N protein 1401 CH ₃ protein 1509 phenylalanine	1014 tryptophan 1156 C-N protein 1236 C-N stretch/ N-H bending of amide III 1548 N-H amide II 1616 C=C tyrosine/tryptophan	1002 phenylalanine 1156 C-N protein 1480 purine bases 1674 C=O amide I

A) CLL (PUP) – CLL (PTP), B) CLL (PUP) – Healthy, C) CLL (PTP) – Healthy

It can be seen that there are higher DNA, nucleic acid and protein contributions in the two CLL pathologies, which corresponds with high cell proliferation. There also appears to be higher lipid and carbohydrate contributions in the healthy data. However there are some discrepancies between the protein content with higher amide I and amide III contributions seen in the CLL data, and higher amide II in the healthy data. Although certain biochemical differences can be identified by this peak assignment, there is still a lot of variation and disagreement between biochemical contributions in the three pathologies. Peak assignment is only speculative as the data obtained from literature is highly varied. Peaks may also represent a variety of biochemical constituents and thus may be masked or shifted due to other components. Baseline, signal strength and noise also varies between the spectra causing any small differences to be exaggerated. All of the sample types and pathologies will display fairly similar and complex biochemistry, with variation surrounding small peak shifts as well as the absorbance/ intensity of certain components being the main source of differentiation. However, these mean spectra appear to indicate that there are not very strong differences between the two CLL pathologies. It is therefore necessary to use multivariate statistical analysis in order to analyse the entire data rather than individual peaks.

5.4. Principal Component Analysis

As described in section 2.3.3 PCA is an unsupervised method, which is used to identify natural separation in the data by reducing the complex spectral information into fewer components that represent the largest biochemical differences. Twenty principal components (PC) were used to produce PC loads, which relate to the spectral component and PC scores, which relate to the abundance of that component in the data. Analysis of variance (ANOVA) was used to identify which of the PC scores were the most significant by looking at the difference in the means of the groups relative to their variance. It calculates an F-value, which corresponds to the ability of the PC score to differentiate between the spectra of the different pathologies. The higher the F-value, the better the separation. The critical value of F (F_{crit}) is then calculated from the F-

distribution (Otto, 2007). F_{crit} is calculated with a 95 % confidence level for the ability of ANOVA to differentiate between the groups (Table 5.12).

Table 5.12. F_{crit} Values of the Seven Sample Groups

Sample Group	F_{crit} Value
Buffy Coat IR	3.72
Blood Cytospin IR	3.71
Blood Droplet IR	3.71
Buffy Coat Raman	3.70
Blood Cytospin Raman	3.72
Blood Droplet Raman	3.70
Plasma Raman	3.71

An F-test was used to identify if the ratio of the variance within groups (unexplained variance) and between groups (explained variance) was significant. This is known as the F-ratio. The F-ratio was plotted for each of the 20 PCs (Figure 5.14). Any PCs with an F-ratio higher than the F_{crit} line (shown in red) are statistically significant and the two PCs with the highest F-ratios were selected to generate PC score plots. The PC loads, which represent the spectral variance, that were chosen by ANOVA are shown in Figure 5.15 with the PC score plots for these loads displayed in Figure 5.17.

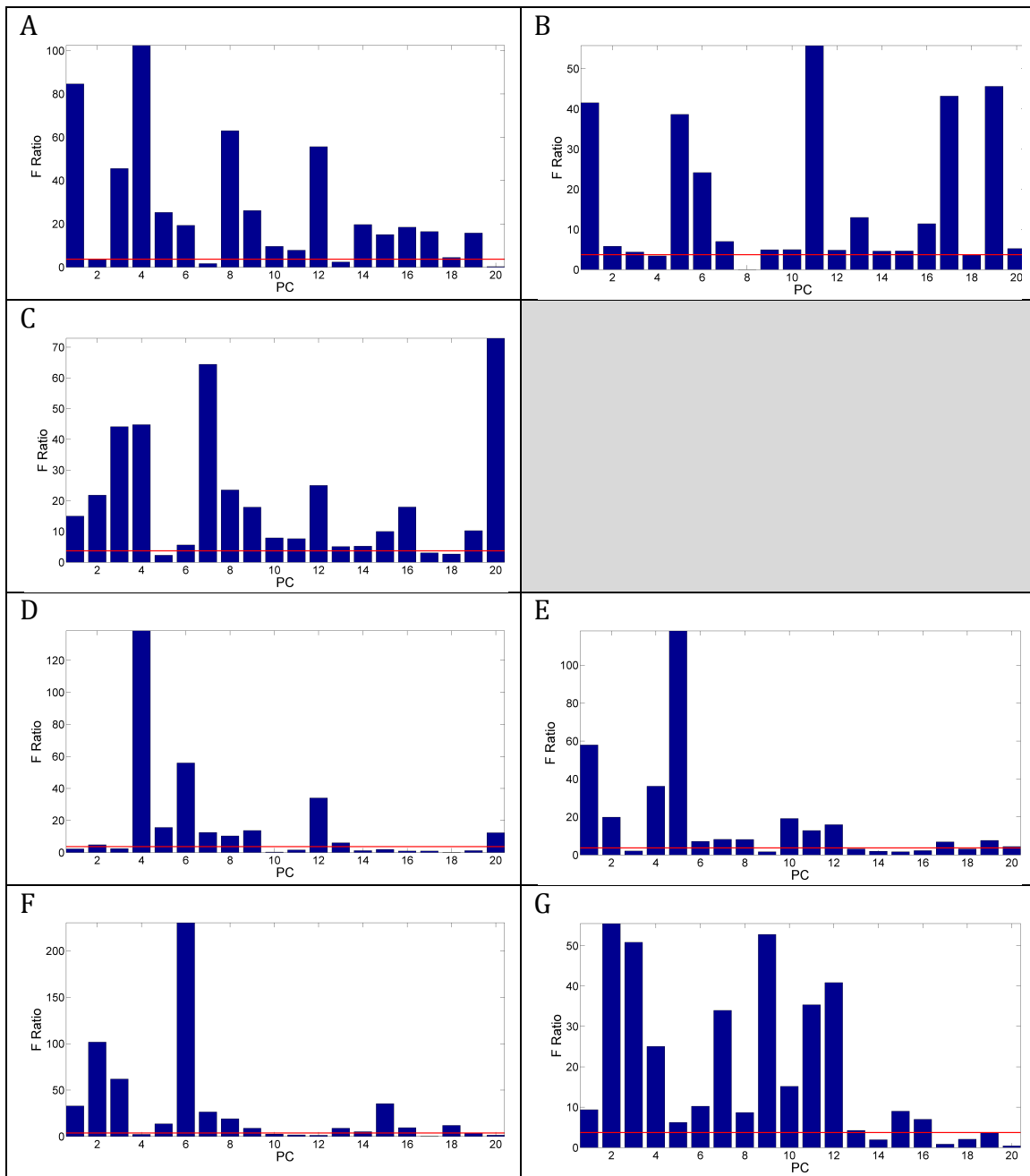
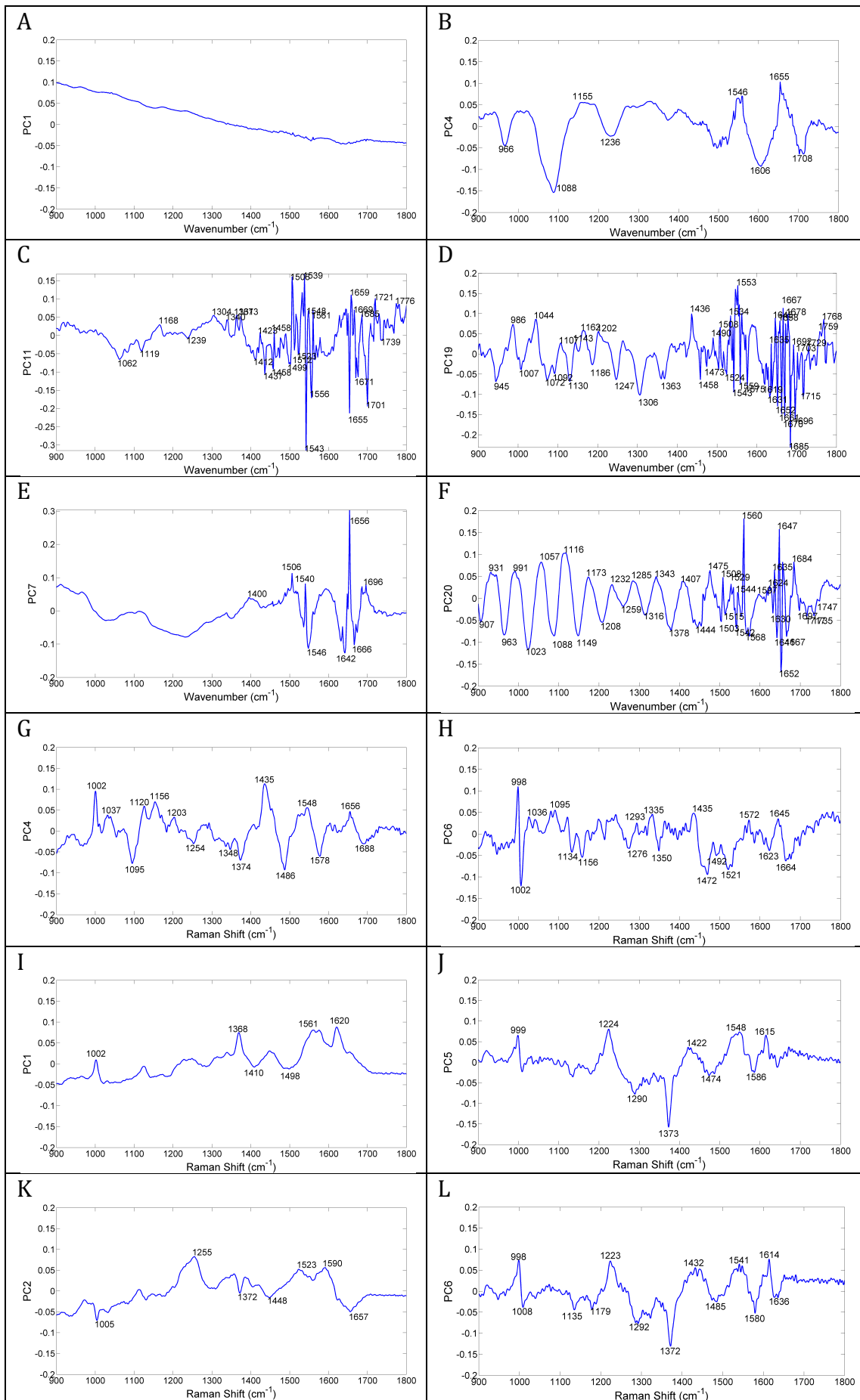


Figure 5.14. Plots of the F-Ratio for Each PC; A) buffy coat IR, B) blood cytospin IR, C) blood droplet IR, D) buffy coat Raman, E) blood cytospin Raman, F) blood droplet Raman, G) plasma Raman (red line indicates F_{crit})



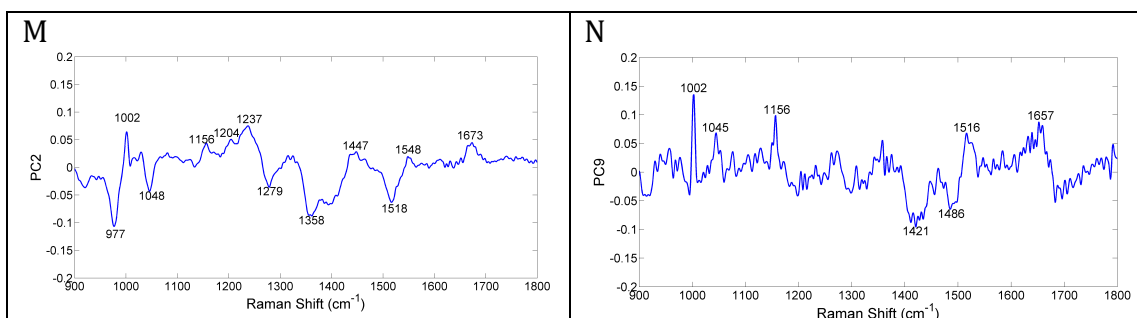


Figure 5.15. PC Loads for Each Dataset as Selected by ANOVA; A) buffy coat IR PC1, B) buffy coat IR PC4, C) blood cytopsin IR PC11, D) blood cytopsin IR PC19, E) blood droplet IR PC7, F) blood droplet IR PC20, G) buffy coat Raman PC4, H) buffy coat Raman PC6, I) blood cytopsin Raman PC1, J) blood cytopsin Raman PC5, K) blood droplet Raman PC2, L) blood droplet Raman PC6, M) plasma Raman PC2, N) plasma Raman PC9

Some of the major peaks observed in the PC loads have been labelled in Figure 5.15. The peaks can be both positive and negative depending on whether the contribution of a biochemical constituent is high or low. The most spectral variance is usually observed in the first few PCs and thus are more desirable as noise is likely to be observed in later PCs, as can be seen in PC19 for IR blood cytopsin (Figure 5.15D) and PC20 for IR blood droplet (Figure 5.15F). However, the variation between samples is larger than the variation between pathology, therefore the first few PC's are not necessarily the most important. Using the buffy coat IR data as an example (Figure 5.16), PCs 1 and 4 have been selected by ANOVA as being the most statistically significant at differentiating between the buffy coat IR spectra of the different pathologies. PC1, which should show the most spectral variance, does not appear to show any biochemical peaks. It is likely that a mixture of artefact, such as a baseline or effects from the environment, and useful biochemical information is the cause. Therefore excluding it might also exclude relevant biochemical information. PC4 however shows several biochemical contributions with particularly large peaks relating to protein (1546 cm^{-1} , 1606 cm^{-1} and 1655 cm^{-1}) and DNA (966 cm^{-1} , 1088 cm^{-1})

contributions, with other peaks relating to carbohydrates (1155 cm^{-1}) and RNA (1708 cm^{-1}).

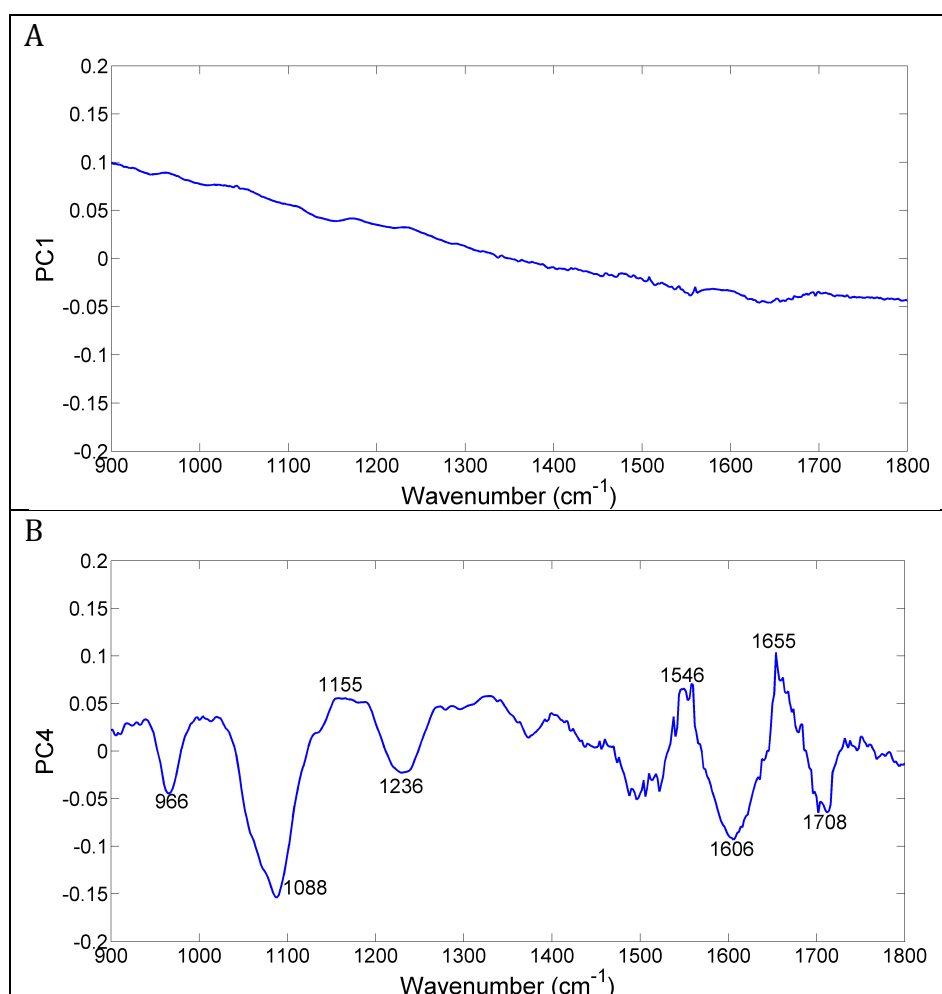
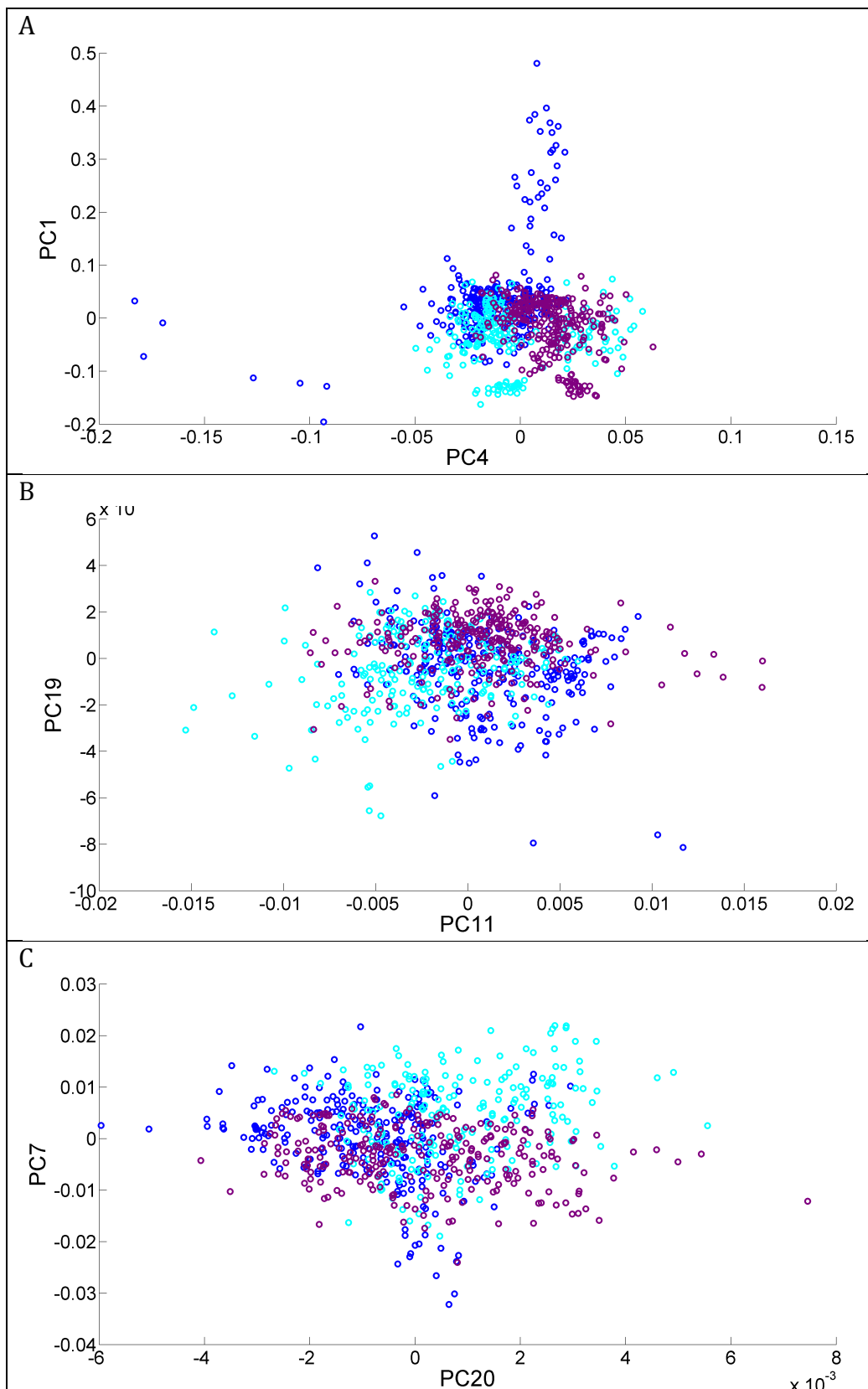
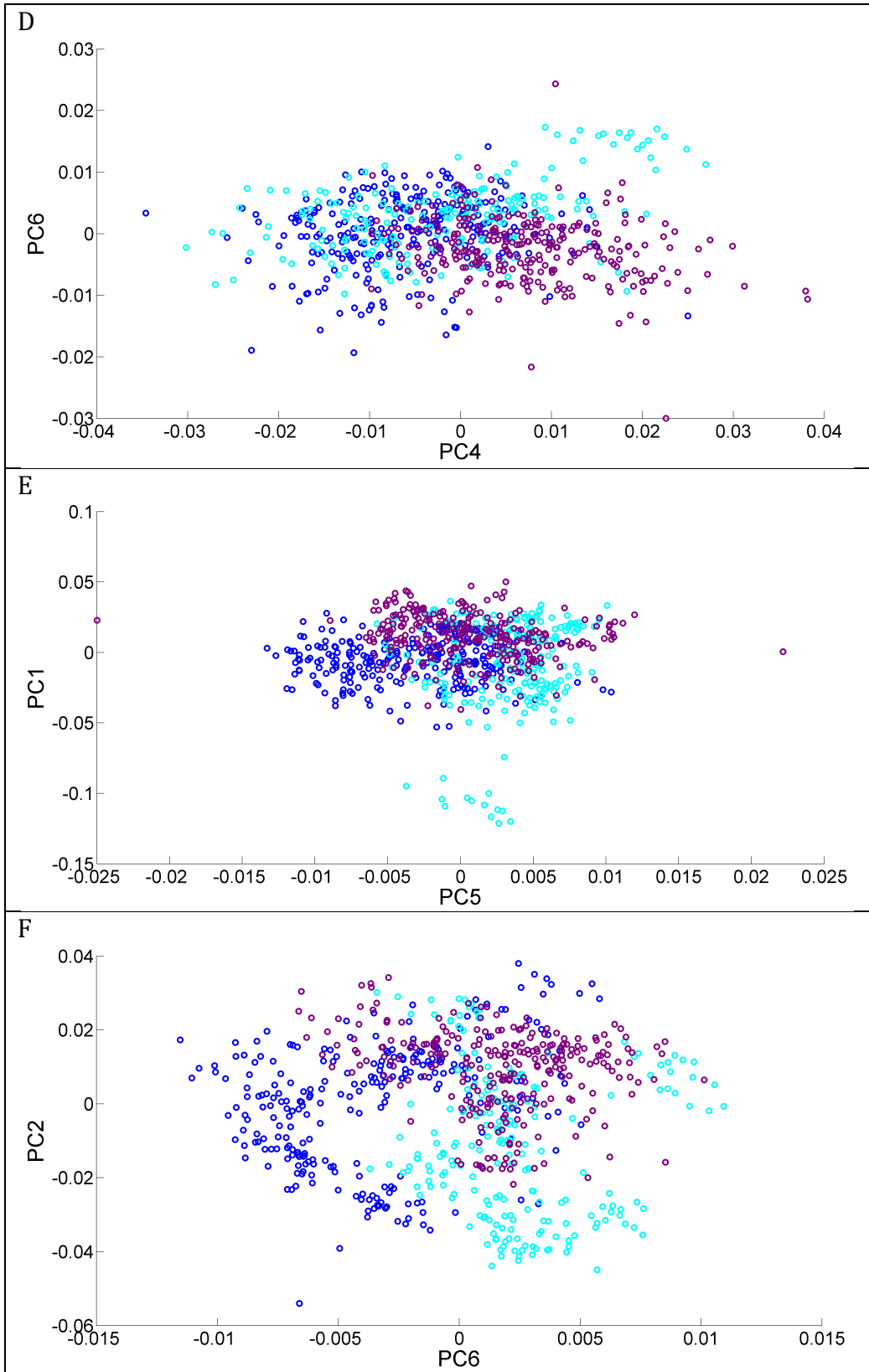


Figure 5.16. PC Loads for IR Buffy Coat; A) PC1, B) PC4

The scores of the two most statistically significant PCs were then plotted against each other in a scatter plot to identify any natural separation that may occur in the data (Figure 5.17).





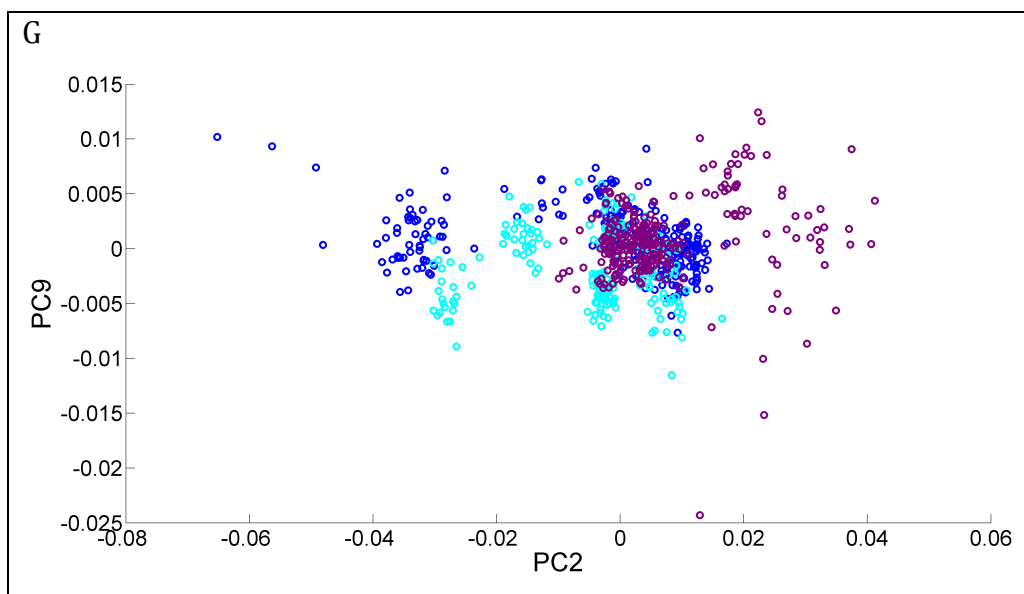


Figure 5.17. PC Score Plots; A) buffy coat IR, B) blood cytospin IR, C) blood droplet IR, D) buffy coat Raman, E) blood cytospin Raman, F) blood droplet Raman, G) plasma Raman, purple: healthy, dark blue: CLL (PUP), light blue: CLL (PTP)

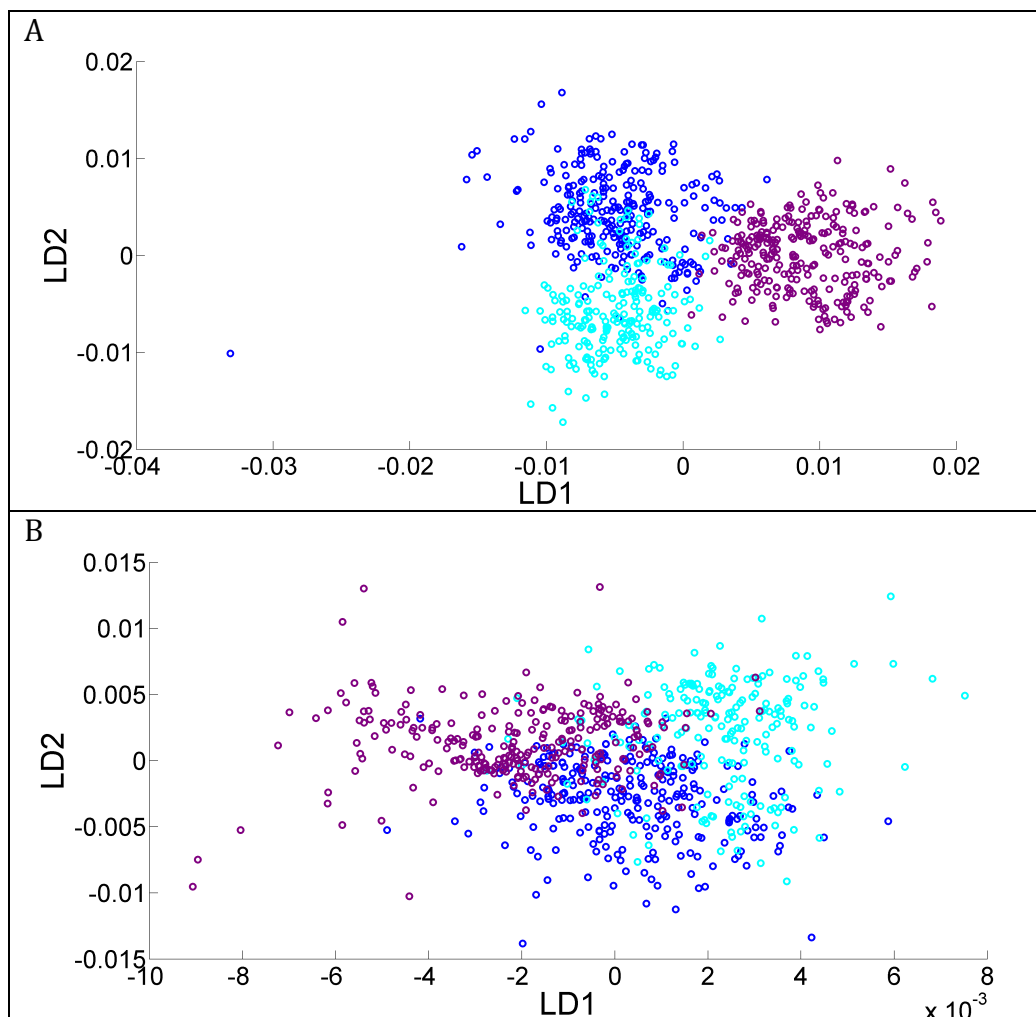
It can be seen in the IR buffy coat PC score plot (Figure 5.17A) that there is some separation between healthy (purple) and the two CLL datasets (light blue and dark blue). The purple data points lie in the positive region of PC4 and thus are separated by amide I, amide II and carbohydrate contributions, as indicated by the positive peaks in the PC load (Figure 5.16). CLL (PUP) (dark blue) and CLL (PTP) (light blue) overlap and trend towards the negative region of PC4 in the scatter plot. These are therefore separated from healthy according to changes in DNA and RNA as shown by the negative peaks in the PC load. Separation can also be seen between the healthy and the two CLL pathologies in the buffy coat Raman plot using PC4 (Figure 5.17D). However there is less obvious separation in the other datasets.

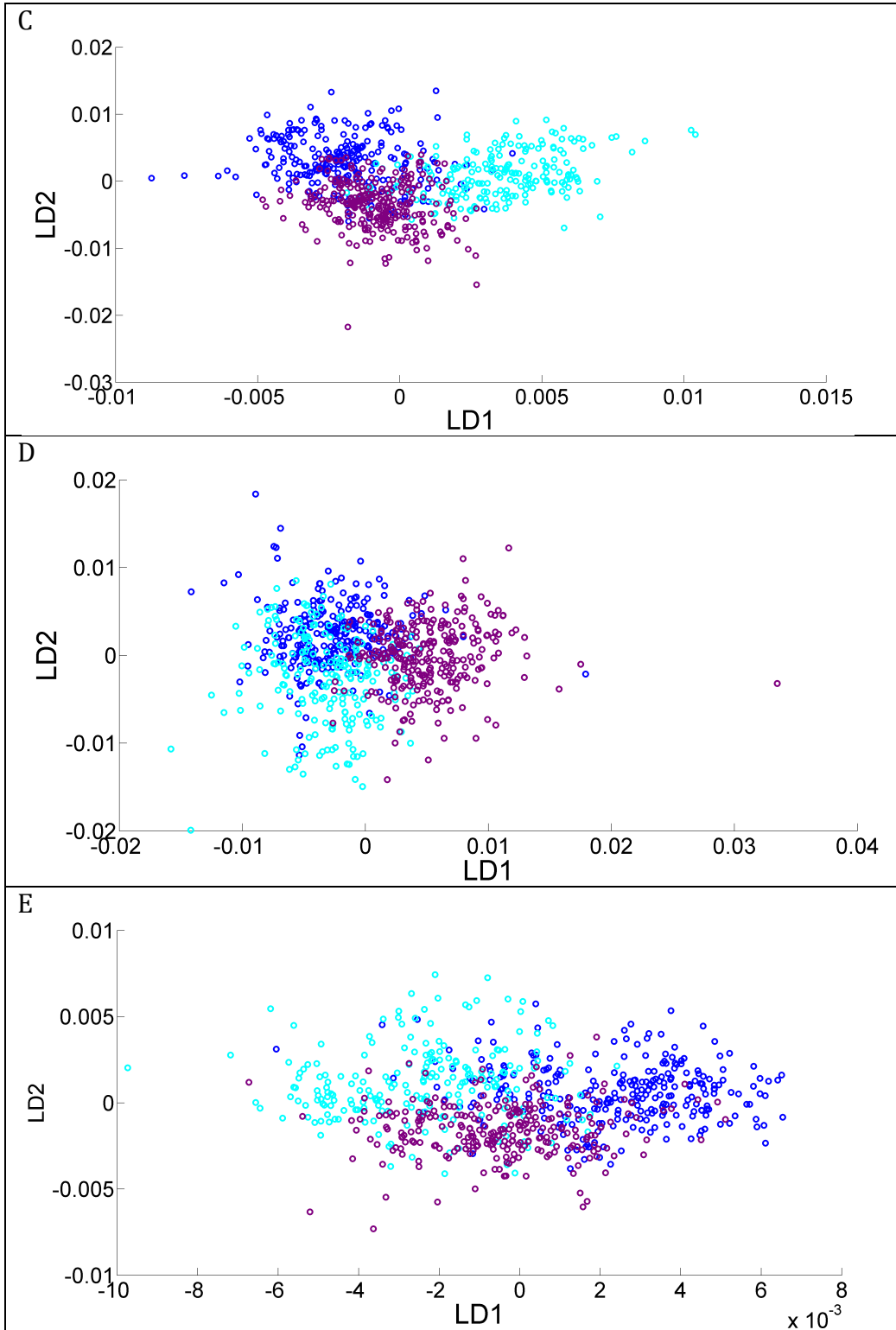
Although these score plots show some natural separation in the data, ANOVA only uses the two most significant PCs to generate scores plots. It can be seen from the plots in Figure 5.14 that more than two PCs contribute to the differences in the data. ANOVA also separates the data based on averages of the

means. If there are any outliers in the data, as can be seen in several of the PC score plots, the mean will be shifted and the ability of the PC to separate the groups will be distorted by these outliers. In order to improve the separation of the pathologies LDA was used.

5.5. Linear Discriminant Analysis

From the F-ratio plots calculated by ANOVA (Figure 5.14), it is clear that more than two PCs have statistically significant contributions to the separation of the data. For many of the sample groups most of the 20 PCs showed high variance, thus using just two did not achieve maximal separation. LDA uses all 20 of the PCs generated in PCA and thus the separation between the groups is maximised according to a linear discriminant function. The LDA scatter plots for the three pathologies in all seven of the datasets are shown in Figure 5.18.





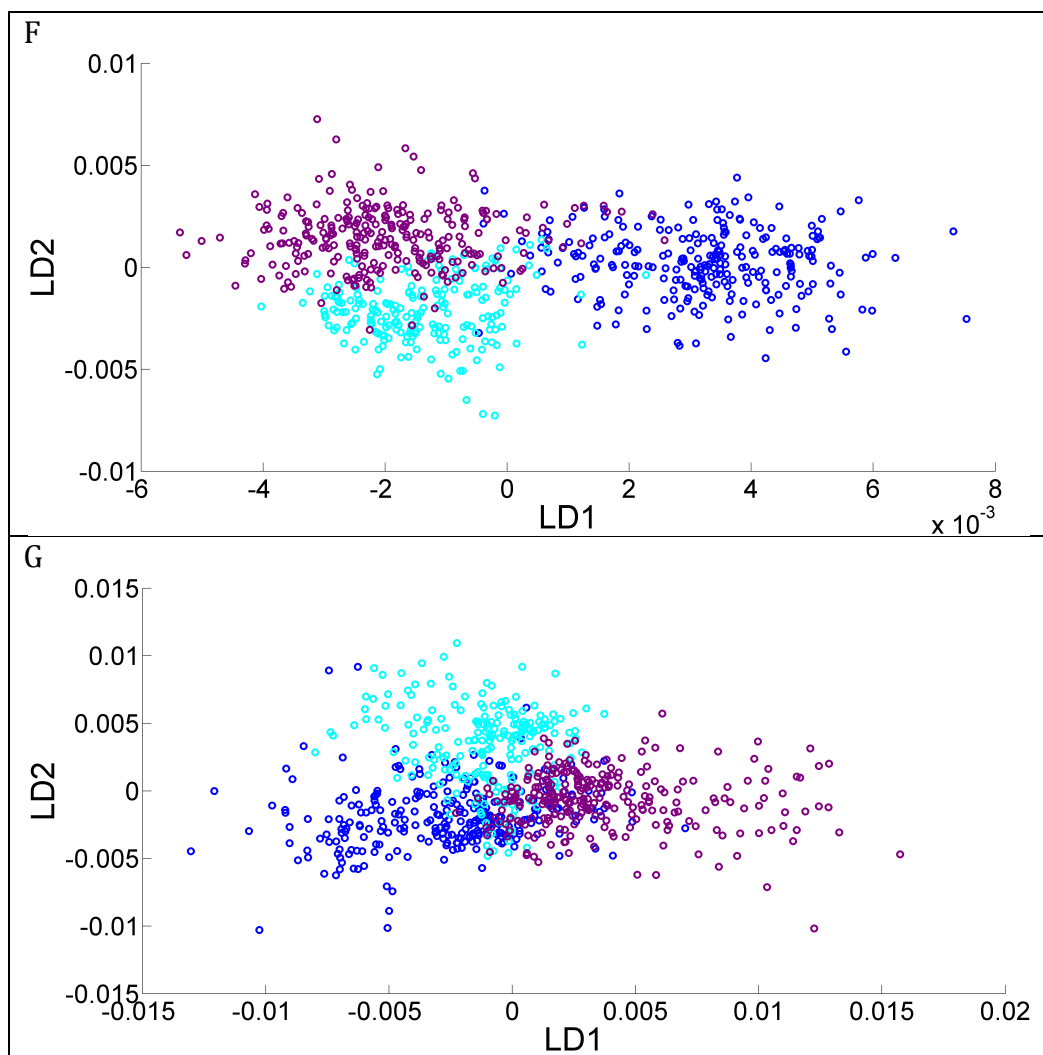


Figure 5.18. LDA Plots; A) buffy coat IR, B) blood cytopsin IR, C) blood droplet IR, D) buffy coat Raman, E) blood cytopsin Raman, F) blood droplet Raman, G) plasma Raman, purple: healthy, dark blue: CLL (PUP), light blue: CLL (PTP)

The separation that can be observed between the three pathologies is greatly improved using LDA. The data for each pathology group are more tightly clustered together and tend to lie further away from the other pathology groups. However, some overlapping of the groups still occurs. This is particularly apparent in the data for the two CLL pathologies thus in order to assess the ability of the PCA followed by LDA classification model, LOSOCV was used.

5.6. Leave One Sample Out Cross Validation

As previously described, the PCA followed by LDA classification model was run using all 20 PCs. A total of 28 healthy and CLL samples were analysed, with a total of 840 spectra collected from each of the sample types in both IR and Raman. In order to evaluate the efficiency of the classification model for each dataset, LOSOCV was run. All of the data from one sample was removed from the model and the analysis rerun using the remaining data to identify if the model could classify the 'left out' samples as accurately. This was done repeatedly with all of the data for each sample. The sensitivities (SN) and specificities (SP) of the classification models are displayed in Tables 5.13 to 5.19

Table 5.13. LOSOCV Results for Buffy Coat IR

		PREDICTED			
		CLL (PUP)	CLL (PTP)	Healthy	TOTAL
	SN %	59	36	77	
	SP %	69	82	87	
TRUE	CLL (PUP)	161	81	30	272
	CLL (PTP)	101	75	34	210
	Healthy	48	15	207	270
					752

59 % of the spectra were correctly classified

Table 5.14. LOSOCV Results for Blood Cytospin IR

		PREDICTED			
		CLL (PUP)	CLL (PTP)	Healthy	TOTAL
	SN %	17	45	19	
	SP %	67	57	65	
TRUE	CLL (PUP)	47	116	107	270
	CLL (PTP)	55	95	59	209
	Healthy	102	114	51	267
					746

26 % of the spectra were correctly classified

Table 5.15. LOSOCV Results for Blood Droplet IR

		PREDICTED			
		CLL (PUP)	CLL (PTP)	Healthy	TOTAL
	SN %	31	36	57	
	SP %	75	87	49	
TRUE	CLL (PUP)	79	34	138	251
	CLL (PTP)	34	75	97	206
	Healthy	83	34	158	276
					733

43 % of the spectra were correctly classified

Table 5.16. LOSOCV Results for Buffy Coat Raman

		PREDICTED			
		CLL (PUP)	CLL (PTP)	Healthy	TOTAL
	SN %	36	20	80	
	SP %	69	74	79	
TRUE	CLL (PUP)	95	119	47	261
	CLL (PTP)	113	42	53	208
	Healthy	33	21	215	269
					738

48 % of the spectra were correctly classified

Table 5.17. LOSOCV Results for Blood Cytospin Raman

		PREDICTED			
		CLL (PUP)	CLL (PTP)	Healthy	TOTAL
	SN %	55	29	49	
	SP %	74	77	66	
TRUE	CLL (PUP)	149	51	71	271
	CLL (PTP)	60	61	92	213
	Healthy	67	71	131	269
					753

45 % of the spectra were correctly classified

Table 5.18. LOSOCV Results for Blood Droplet Raman

		PREDICTED			
		CLL (PUP)	CLL (PTP)	Healthy	TOTAL
	SN %	70	48	66	
	SP %	89	80	74	
TRUE	CLL (PUP)	178	46	31	255
	CLL (PTP)	19	98	88	205
	Healthy	32	59	179	270
					730

62 % of the spectra were correctly classified

Table 5.19. LOSOCV Results for Plasma Raman

		PREDICTED			
		CLL (PUP)	CLL (PTP)	Healthy	TOTAL
	SN %	49	41	64	
	SP %	71	78	80	
TRUE	CLL (PUP)	137	84	57	278
	CLL (PTP)	84	86	42	212
	Healthy	58	39	173	270
					760

52 % of the spectra were correctly classified

Overall the sensitivities and specificities of the different sample types are low, particularly between the two CLL pathologies. The poor performance of the classification model is also identified by the low percentage of spectra correctly classified across all datasets, as highlighted in Figure 5.19. The low number of samples collected and hence number of spectra obtained may have contributed to this, as the effect of random outliers will be enhanced in a smaller dataset.

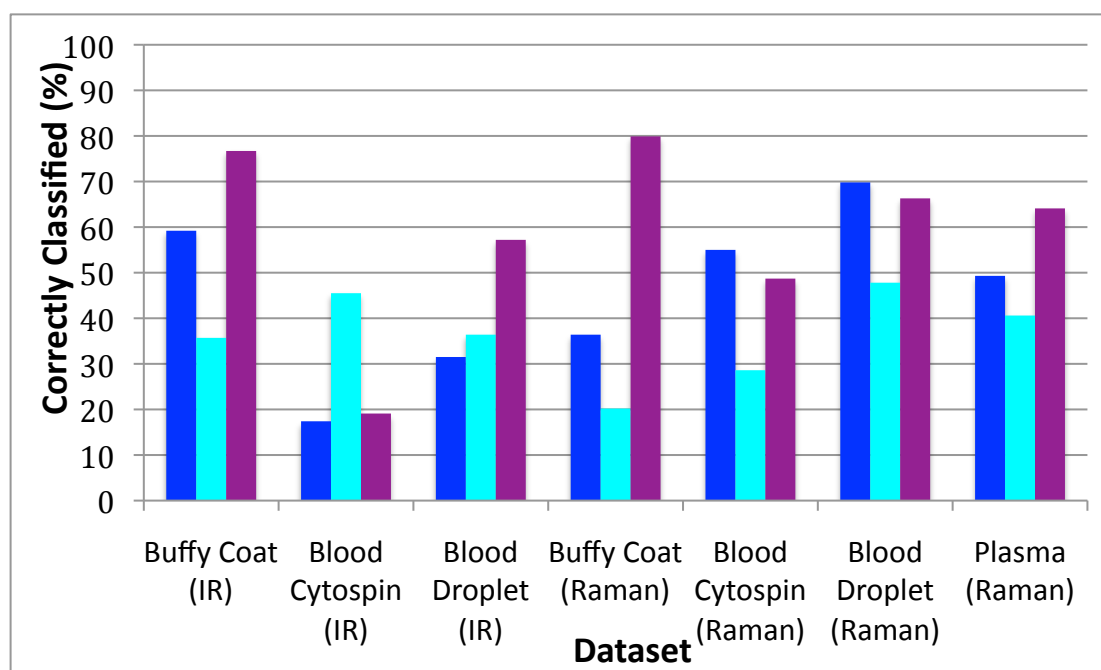


Figure 5.19. Percentage of Spectra In Each Dataset Correctly Classified by LOSOCV; dark blue: CLL (PUP), light blue: CLL (PTP), purple: (Healthy)

The blood droplet Raman and buffy coat IR methods show better separation of the two individual CLL pathologies compared to the other methods. This can be seen by the higher sensitivities and specificities, however classification is still poor with less than 50 % of the CLL (PTP) correctly classified. Throughout it has been clear that there is poor separation between the two CLL populations with patients who had previously been treated for CLL and patients who were newly diagnosed showing little biochemical variance. Because of this the classification model was rerun, with the two CLL populations grouped together.

5.7. Two Group Model

All data from the two CLL pathologies were combined and the classification model rerun as before. The sensitivities and specificities of the two group classification models are displayed in Tables 5.20 to 5.26.

Table 5.20. Two Group LOSOCV Results for Buffy Coat IR

		PREDICTED		
		CLL	Healthy	TOTAL
	SN %	83	76	
	SP %	76	83	
TRUE	CLL	398	85	483
	Healthy	64	206	290
				773

78 % of the spectra were correctly classified

Table 5.21. Two Group LOSOCV Results for Blood Cytospin IR

		PREDICTED		
		CLL	Healthy	TOTAL
	SN %	54	28	
	SP %	28	54	
TRUE	CLL	259	220	479
	Healthy	191	76	267
				746

45 % of the spectra were correctly classified

Table 5.22. Two Group LOSOCV Results for Blood Droplet IR

		PREDICTED		
		CLL	Healthy	TOTAL
	SN %	45	60	
	SP %	60	45	
TRUE	CLL	206	251	457
	Healthy	107	163	270
				727

51 % of the spectra were correctly classified

Table 5.23. Two Group LOSOCV Results for Buffy Coat Raman

		PREDICTED		
		CLL	Healthy	TOTAL
	SN %	77	82	
	SP %	82	77	
TRUE	CLL	359	110	469
	Healthy	49	220	269
				738

78 % of the spectra were correctly classified

Table 5.24. Two Group LOSOCV Results for Blood Cytospin Raman

		PREDICTED		
		CLL	Healthy	TOTAL
	SN %	54	54	
	SP %	54	54	
TRUE	CLL	260	224	484
	Healthy	124	145	269
				753

54 % of the spectra were correctly classified

Table 5.25. Two Group LOSOCV Results for Blood Droplet Raman

		PREDICTED		
		CLL	Healthy	TOTAL
	SN %	69	76	
	SP %	76	69	
TRUE	CLL	316	144	460
	Healthy	66	204	270
				730

71 % of the spectra were correctly classified

Table 5.26. Two Group LOSOCV Results for Plasma Raman

		PREDICTED		
		CLL	Healthy	TOTAL
	SN %	69	66	
	SP %	66	69	
TRUE	CLL	335	155	490
	Healthy	93	177	270
				760

67 % of the spectra were correctly classified

In all of the datasets the sensitivities, specificities and number of spectra correctly classified increased significantly in the two group model compared to the previously run three group model. The number of CLL spectra correctly identified is also much higher, with the buffy coat IR showing 83 %, buffy coat Raman 77 % and blood droplet Raman achieving 69 % correct classification (Figure 5.20). The performance of the seven datasets in both the two and three group models are listed in Table 5.27 in order of ability to identify spectra correctly.

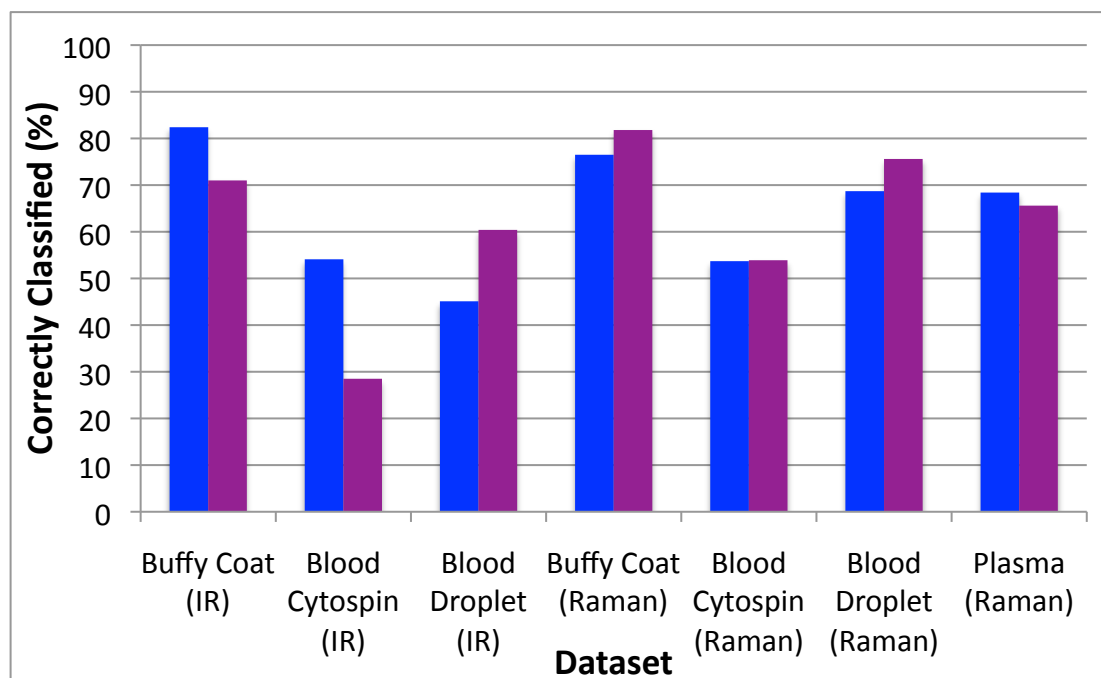


Figure 5.20. Percentage of Spectra In Each Dataset In the Two Group Model Correctly Classified by LOSOCV; dark blue: CLL, purple: healthy

Table 5.27. Overall Performance of Datasets in the Two and Three Group Models

Model	Dataset	% All Data Correctly Classified	% SN Range	% SP Range
2 Group	Buffy Coat (IR)	78	76-83	76-83
2 Group	Buffy Coat (Raman)	78	77-82	77-82
2 Group	Blood Droplet (Raman)	71	69-76	69-76
2 Group	Plasma (Raman)	67	66-69	66-69
3 Group	Blood Droplet (Raman)	62	48-70	74-89
3 Group	Buffy Coat (IR)	59	36-77	69-87
2 Group	Blood Cytospin (Raman)	54	54	54
3 Group	Plasma (Raman)	52	41-64	71-80
2 Group	Blood Droplet (IR)	51	45-60	45-60
3 Group	Buffy Coat (Raman)	48	20-80	69-79
2 Group	Blood Cytospin (IR)	45	28-54	28-54
3 Group	Blood Cytospin (Raman)	45	29-55	66-77
3 Group	Blood Droplet (IR)	43	31-57	49-87
3 Group	Blood Cytospin (IR)	26	17-45	57-67

Black: two group model, blue: three group model

In order to try to improve the classification model further by reducing the influence of noise, a more conservative PCA followed by LDA classification model was run using only 12PCs. An overview of the results of the LOSOCV using 12 PCs are shown in Tables 5.28 and 5.29 with a comparison of the number of spectra correctly classified in the complete (20 PC) and conservative (12 PC) models displayed in Table 5.30.

Table 5.28. FTIR LOSOCV Results Using 12PCs

	Buffy Coat		Blood Cytospin		Blood Droplet	
	SN%	SP%	SN%	SP%	SN%	SP%
CLL	78	93	52	18	58	66
Healthy	93	78	18	52	66	58

Table 5.29. Raman LOSOCV Results Using 12PCs

	Buffy Coat		Blood Cytospin		Blood Droplet		Plasma Cytospin	
	SN%	SP%	SN%	SP%	SN%	SP%	SN%	SP%
CLL	77	80	54	54	65	81	65	68
Healthy	80	77	54	54	81	65	68	65

Table 5.30 Comparison of the Performance of the Complete and Conservative Two Group Models

Dataset	% Correctly Classified (Conservative Model)	% Correctly Classified (Complete Model)
Buffy Coat (IR)	83	78
Buffy Coat (Raman)	78	78
Blood Droplet (Raman)	71	71
Plasma (Raman)	66	67
Blood Droplet (IR)	61	51
Blood Cytospin (Raman)	54	54
Blood Cytospin (IR)	40	45

In the conservative model the number of correctly classified spectra improved for the buffy coat IR and blood droplet Raman datasets and remained the same for the buffy coat Raman, blood droplet Raman and blood cytospin Raman datasets. The plasma Raman and blood cytospin IR datasets were actually better at correctly classifying the data using all 20 PCs. However their performance is still quite poor compared to the other datasets.

5.8. Discussion

The difference spectra showed that significant spectral differences existed between the three pathology groups, particularly between the CLL pathologies and the healthy group. The major differences seen between the CLL and healthy groups being attributed to biochemical contributions from proteins and DNA. PCA identified some natural separation between the two CLL pathologies and the healthy group, which was further separated using all 20 PCs in LDA. LOSOCV showed that the initial three group model was unable to differentiate between the two CLL pathologies very well. When these pathology groups were combined in the two group model, the classification of CLL from healthy was greatly improved across all sample types, identifying that misclassification due to a patient having previously been treated for CLL is not very likely. In order to try to improve the classification model further by reducing the influence of noise, the two group model was rerun using 12 PCs. This moderately improved the performance of some of the datasets suggesting that beyond 12 PCs noise is more likely to be contributing to the variance than biochemical differences.

It must be highlighted that in the previous study by Babrah (2009), 99.5 % correct classification of the buffy coat IR model was achieved. Here only 83 % correct classification was achieved in the conservative two group model. Although this is significantly lower, in the previous study 1438 spectra were acquired from ten blood samples, an average of 140 spectra per sample, whereas here 840 spectra were acquired from 28 samples, 30 spectra from each. As the performance of the blood droplet Raman model was very similar to the buffy coat IR in this study, obtaining a higher number of samples and spectra may improve the model. This could be tested by rerunning the models and leaving two or three samples out each time. It was also identified from these results that cytopinning whole blood samples onto slides was not a very good method choice for CLL diagnosis. Due to the sample being a heterogeneous layer of cells, the cellular content across the slide was highly varied whereas the blood droplets created a homogenous 'soup' of biochemical content from which more consistent spectral collection could be obtained across the sample.

6. RESULTS AND DISCUSSION: LYMPHOMA ANALYSIS

Due to the low number of samples obtained, this chapter on lymphoma is presented as a feasibility study in order to assess the methods of sample preparation and spectroscopic analysis that were used. Five fine needle aspirate biopsies (FNAB) were obtained. The first sample that was collected was subject to both cytopsin centrifugation and drop coating deposition spectroscopy (DCDS) measured by Raman and Fourier Transform Infrared (FTIR) as described in section 3.5. The remaining samples were only cytopsin due to the interference of phosphate buffered saline (PBS) observed in the white light images and potentially in the data. Unfortunately a record of the pathologies was not maintained.

6.1. Physical Properties

The first lymph node aspirate biopsy (FNAB) received was provided in an unknown volume of phosphate buffered saline (PBS). 150 μl of the sample was cytopsin onto a CaF_2 slide, whilst 1.5 μl was pipetted. The white light images of both samples can be seen in Figure 6.1.

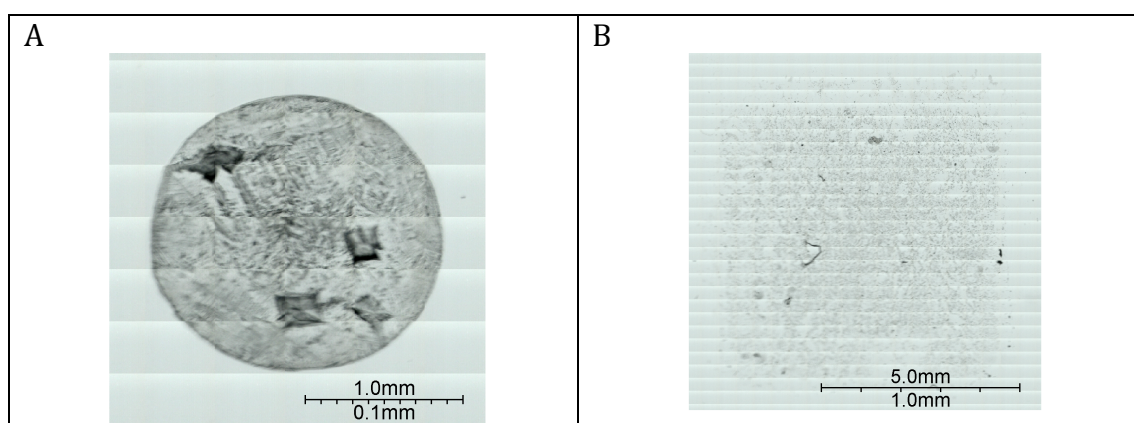


Figure 6.1. White Light Images of the First FNAB Received; A) DCDS, B) Cytospin

As identified in the fibrinogen and globulin droplets (chapter 4), the use of PBS introduces diamond shapes of highly concentrated sample rather than a ring as required for the DCDS method. Ferning patterns can also be seen, which if more samples had been received and further analysis undertaken may have provided further information for diagnosis as suggested by Pearce & Tomlinson (2000) who used these ferning patterns to aid in the diagnosis of ocular abnormalities. The cytopsin sample (Figure 6.1B) is very thin. A monolayer of cells should be produced by this method, however as the volume of PBS used for the FNAB was unknown the volume may have been too high for the number of cells obtained. Following this, tubes with known volumes of PBS were provided to clinicians for further sample collections. Samples 2 and 3 contained 2 ml and samples 3 and 4 contained 3 ml PBS.

Figure 6.2 shows the white light images for the remaining four samples. As the volume of PBS in the sample is increased the cytopsin sample becomes thinner on the slide and more difficult to analyse spectroscopically. Thus it was more difficult to guarantee that a spectrum was obtained from cells of interest as opposed to the CaF₂ slide. A volume of PBS no greater than 2 ml therefore seems appropriate. The results observed for Raman and FTIR are displayed in section 6.2.

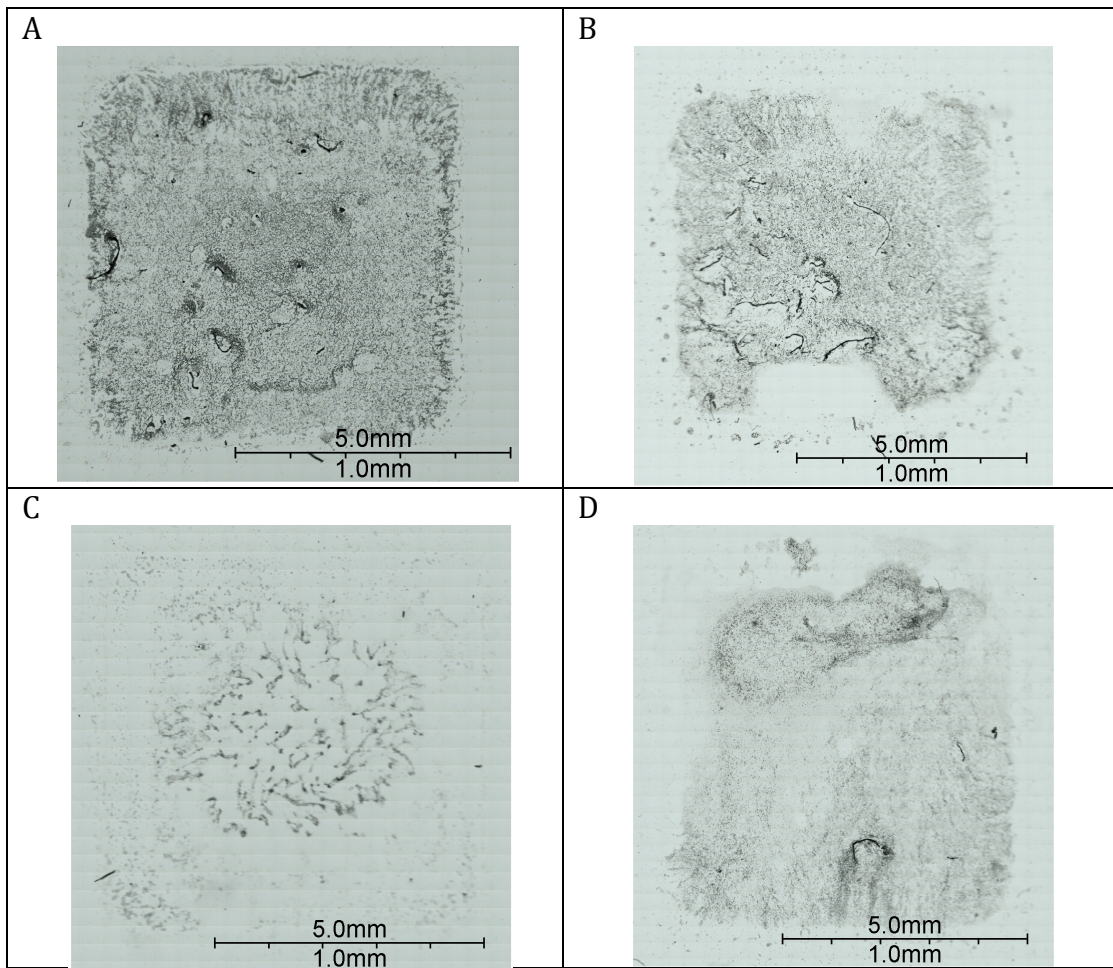


Figure 6.2. Cytospin White Light Images; A) sample 2 in 2ml PBS, B) sample 3 in 2ml PBS, C) sample 4 in 3ml PBS, D) sample 5 in 3ml PBS

6.2. Spectral Analysis of Lymphoma Samples

Thirty Raman and IR point spectra were obtained from each sample using the parameters described in section 3.4.4.3. Mean spectra were plotted and peaks identified using an in house Matlab tool. Figure 6.3 displays the mean Raman and infrared (IR) spectra obtained from the first sample collected.

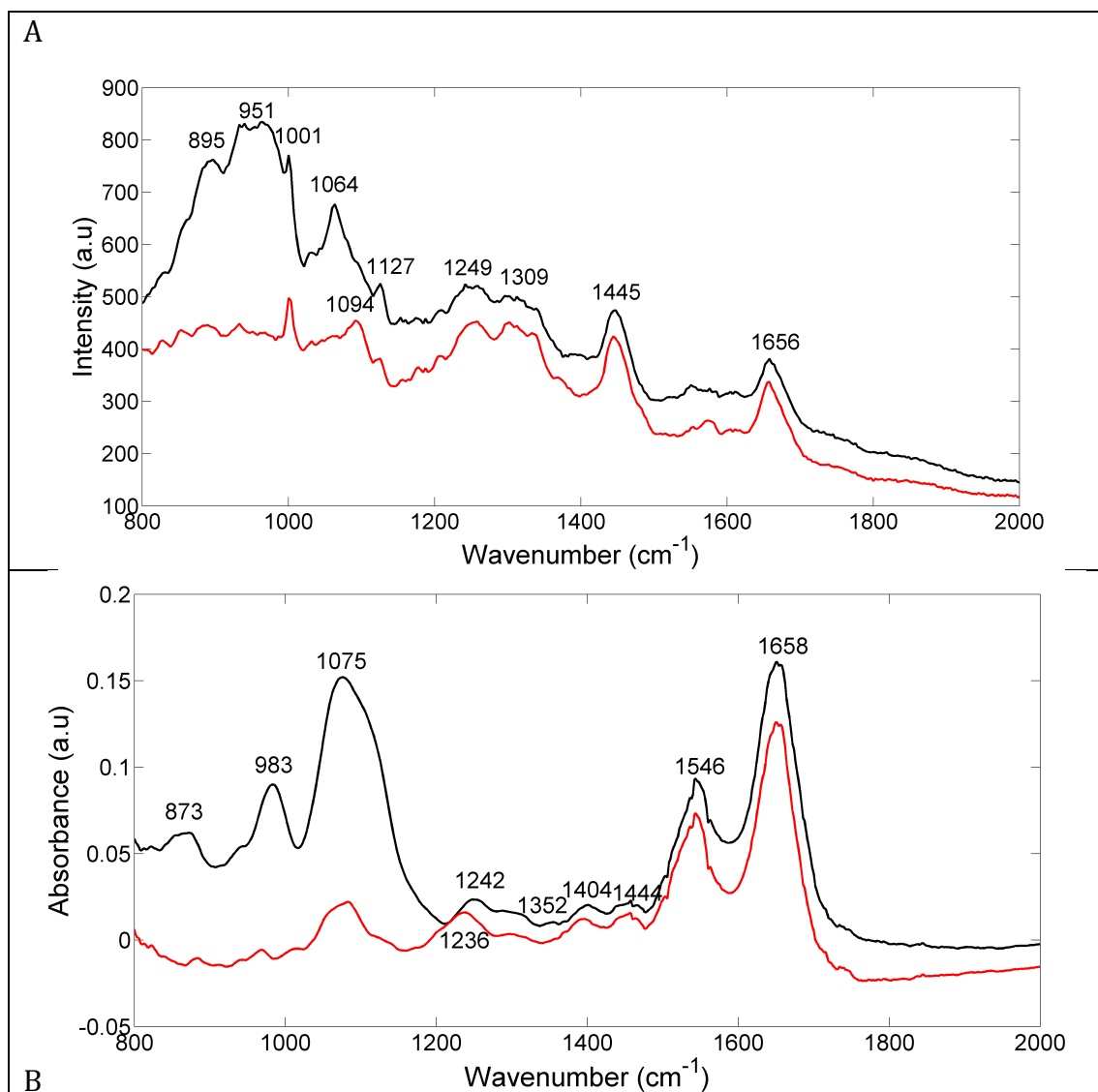


Figure 6.3. Mean Spectra of Sample 1; A) Raman, B) IR, red: cytospin sample, black: DCDS sample

Both the Raman and IR spectra are of good quality, with important biochemical peaks observed. Peaks representing proteins can be seen at 1658 cm⁻¹ for amide I (Chiriboga, 1998a & Mahadevan-Jansen, 1996), 1546 cm⁻¹ for amide II (Liu, 2002) and 1242 cm⁻¹ for amide III (Liu, 2002 & Stone, 2002). Lipid peaks can also be seen between 1320-1460 cm⁻¹ due to the CH₂ and CH₃ vibrations from their long hydrocarbon chains (Liu, 1992 & Nevilliappan, 2002). However peaks from nucleic acids appear to be masked by some very large peaks, particularly in the DCDS spectra. This may be due to high DNA contributions, or it may be due to

interference from other artefacts. Without knowing more about the pathologies of these samples the cause of this high signal cannot be determined from these results alone. As PBS was added to every sample it may have been a source of interference and thus was measured to see if there were any contributions to the spectra (Figure 6.4).

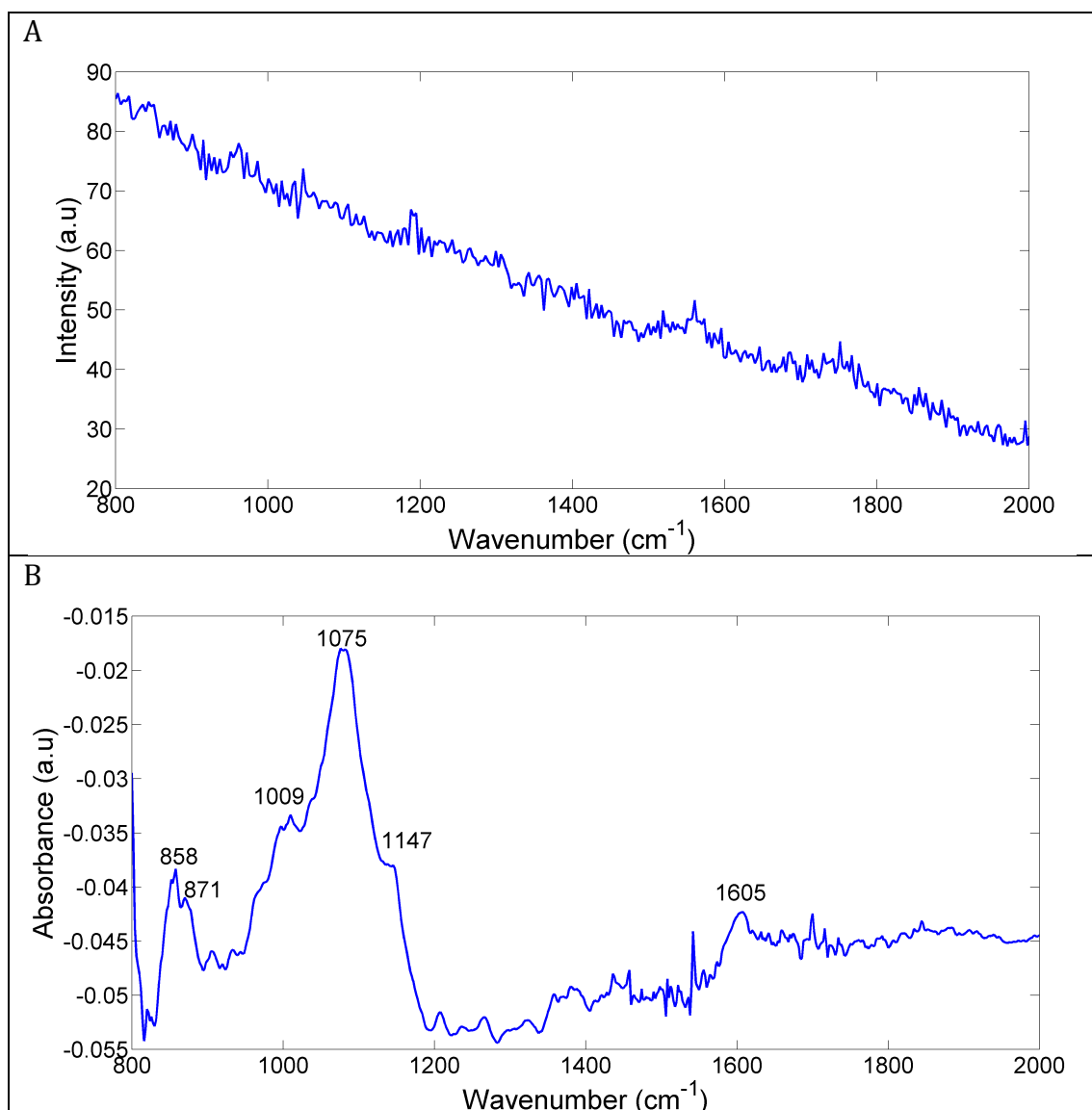


Figure 6.4. PBS Spectra; A) Raman, B) FTIR

PBS does not appear to be Raman active but there are some contributions from PBS to the IR spectra, particularly between 900-1100 cm^{-1} . It may be that due to the high concentrating effect of DCDS, PBS has a more intense signal that is otherwise hidden in the cytospin samples by the presence of other biochemical contributions. However, further work is required in order to confirm this hypothesis. These signals are also present in some of the other four cytospin samples obtained (Figure 6.5). What is interesting about these spectra is how different they all are. This may be attributable to the different cell types or pathologies in the samples but as the pathologies were unknown it is difficult to understand this further. Sample 5 (purple) shows some particularly large peaks in the 900-1100 cm^{-1} region in both Raman and IR. Sample 4 (orange) also has some large peaks in IR but not in Raman. These two samples were diluted in a larger volume of PBS, 3ml instead of 2ml. Again, without knowing more about the samples, whether they are cancerous or benign it is difficult to understand the cause of the peaks in this region.

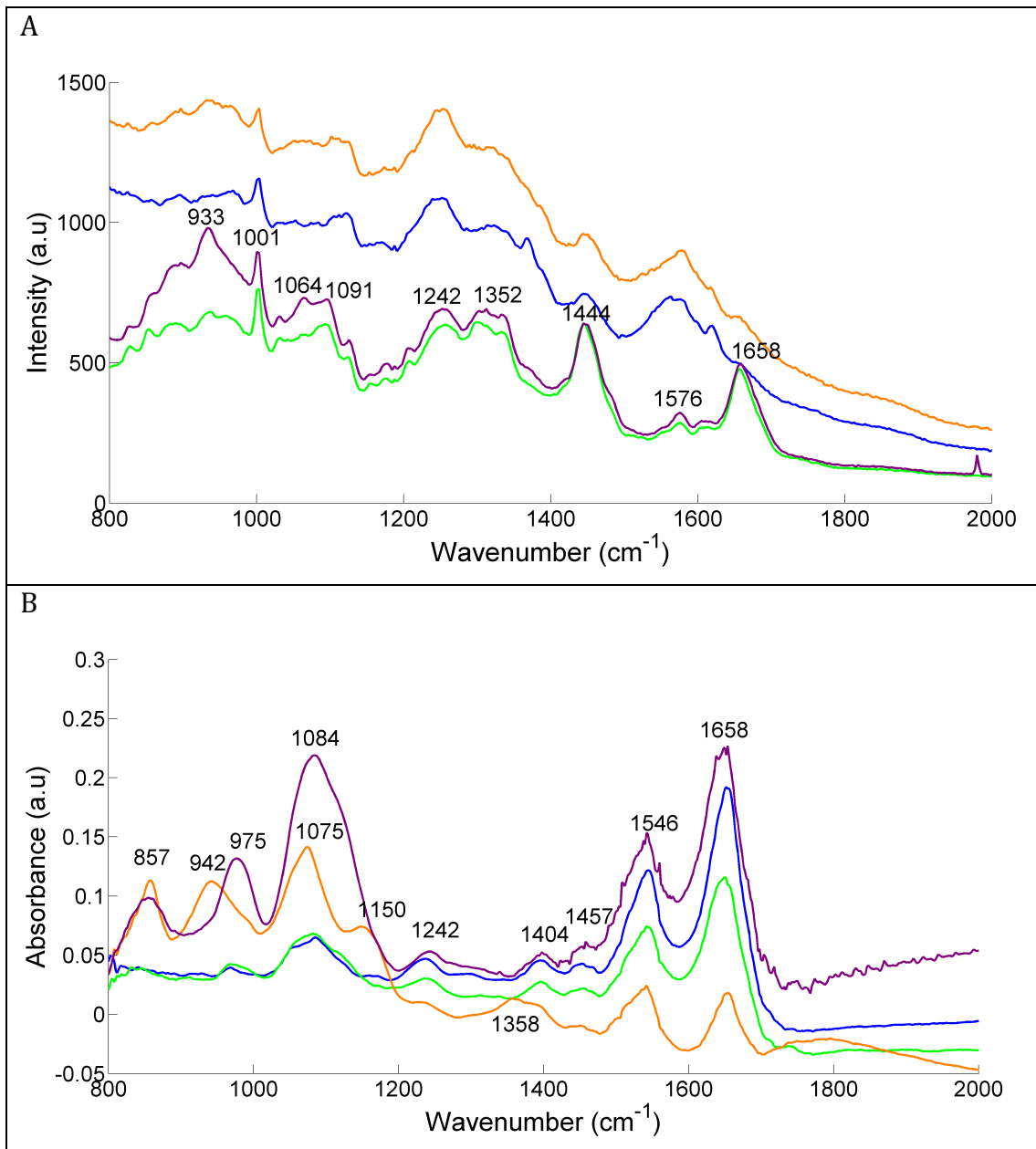


Figure 6.5. Mean Spectra of FNAB Samples; A) Raman, B) IR, blue: sample 2, green: sample 3, orange: sample 4, purple: sample 5.

6.3. Discussion

The spectra obtained show the potential for lymph node aspirate biopsies to be analysed by both Raman and IR spectroscopy. Unfortunately, due to difficulties in obtaining samples and the limited number obtained with unknown pathologies, any further statistical analysis was not appropriate.

FNABs are routinely collected and flushed in PBS in the hospital. Although at this stage it is only speculation that there is spectral interference from PBS, the physical structure of the droplets is poor. As discovered by the blood protein analysis in chapter 4, the formation of diamond shapes is not desirable for good spectral measurements. It would therefore be interesting to see if the samples could be obtained in water for DCDS.

7. CONCLUSIONS

With the aid of multivariate statistical analysis, the aim of this thesis was to investigate the application of Raman and Fourier Transform infrared (FTIR) spectroscopy as a rapid diagnostic tool for leukaemia and lymphoma. The main focus of this thesis was to compare the performance of the different blood fractions prepared by two different techniques; cytopsin centrifugation and Drop Coating Deposition Spectroscopy (DCDS), in diagnosing leukaemia. The DCDS method was explored owing to the low sample volumes and minimal sample preparation required and thus is advantageous over the more laborious cytopsin centrifugation method. Unfortunately the investigation of FTIR and Raman for lymphoma diagnosis did not get very far due to the small number of undiagnosed fine needle aspirate biopsy samples obtained.

7.1. DCDS Method

The results from the preliminary DCDS study revealed it to be a powerful technique for both infrared (IR) and Raman spectroscopic analysis of low concentration protein solutions. It was identified that in order to achieve qualitative biochemical information from the droplets, initial morphological analysis of the formed rings was required in order to avoid taking spectral measurements from areas in the ring where cracks had formed. It was also identified that scattering occurred at the very edges of the rings to a certain extent and by taking spectral measurements a specified distance from the edge these effects could be eliminated from the results. This information was used to develop a technique for leukaemia diagnosis using whole blood samples.

7.2. Leukaemia Classification Model

Morphological analysis allowed the best ring dilution to be measured for each spectroscopic method. Each blood sample obtained had a very different consistency, varying in number of cells, colour and viscosity and thus diluting

blood to obtain a specific concentration or appearance in order to obtain a precise droplet on the slide was extremely difficult. It may have been beneficial to identify the concentrations of some of the analytes in the samples for calibration purposes. However at this early stage it was difficult to know what analytes may have been different in the cancer and healthy samples. A full blood count could have been done to identify the indices and aid in the dilution of the samples for drop coating deposition. However, not only is it time consuming but the white cell count is very high in leukaemia, which causes a reduction in circulating red blood cells thus making comparisons between healthy and cancerous samples quite difficult. Making three dilutions of each blood sample and visually assessing the quality of the droplets was the easiest and most practical option, especially if the technique was to be brought into a clinical environment. The accuracy of this visual technique was confirmed in chapter 4 using albumin droplets, which aided in the determination of acceptable criteria for this process.

A numerical screening method aided in the semi-automated removal of IR data that was saturated. Although some poor quality samples can be removed visually, using a numerical method allows a more accurate technique of eliminating spectra that has interference from background noise or from a sample being too thick for the light to pass through. Although this meant that several plasma samples were removed, the spectra in Figure 5.3 show that no extra biochemical information was obtained from plasma than was identified in the other blood fractions.

The principal component analysis (PCA) followed by linear discriminant analysis (LDA) classification method performed best in the two group model using 12 PCs for both the buffy coat infrared and blood droplet Raman datasets. This was shown by the percentage of correctly classified spectra as calculated by leave one sample out cross validation (LOSOCV). The performance of the classification model for the buffy coat data was not as good at correctly classifying the spectra as shown by Babrah (2009), with 83 % correctly classified compared to 99.5 %.

However, the number of spectra per sample obtained here was much lower, with hundreds rather than thousands of point spectra analysed.

The method of extracting the buffy coat layer is laborious and thus the DCDS method using whole blood was compared in order to minimise sample preparation and overall analysis time. The DCDS preparation in theory provides a homogenous ring signal so only a few measurements need to be made that represent the signal from the blood collected. From these results it would seem that the blood droplet Raman classification model, which achieved 71 % correct classification, was not quite as successful as the buffy coat infrared model, which correctly classified 83 % of the spectra, at diagnosing leukaemia. With the collection and analysis of more samples and further development of the statistical model used, the DCDS method has potential to be improved further, increasing the sensitivity and specificity of the technique in diagnosing cancer. The advantage of using this technique over buffy coat infrared is the smaller sample volumes and minimal sample preparation required. The samples can also be measured almost instantly with Raman spectroscopy as water does not cause interference and thus time could be saved not having to dry the samples first. It may be possible that droplets could be measured whilst still wet, although this could be a problem with low concentration samples. It would be interesting to test this theory out. However the effects of hydration on the biochemistry of samples would also need to be addressed in order to avoid variation in hydration being mistaken for biochemical differences due to cancer.

All patients used for this study were newly diagnosed cases of Chronic Lymphoblastic leukaemia (CLL). Both populations of CLL, previously unseen patients (PUP) and previously treated patients (PTP) showed little biochemical difference and thus when the classification models were rerun with these two groups combined, the overall performances improved. It is therefore unlikely that misclassification due to a patient having previously been treated for CLL would occur.

7.3. Lymphoma

Insufficient samples were collected for appropriate conclusions to be drawn regarding the use of vibrational spectroscopies for the diagnosis of lymphoma. It has been shown that spectra can be taken from fine needle aspirate biopsies and from lymph nodes and thus further work is required into the method of sample preparation, identifying the biochemical differences between normal and lymphoma samples and developing a classification model for successful diagnosis.

7.4. Clinical Applications

Although vibrational spectroscopies, particularly FTIR, have been used to monitor leukaemias and identify the effect of chemotherapy treatment, there is a lack of evidence in the literature that different sub-types of leukaemia can be diagnosed by spectroscopic analysis. It has been identified in this study that CLL can be differentiated from healthy samples using buffy coat IR, buffy coat Raman and blood droplet Raman classification models. Unfortunately due to the number of samples collected it was not possible to look at other leukaemia pathologies. If these could also be diagnosed with the same level of sensitivity and specificity, these techniques have the potential to be applied to clinical practice.

Having worked in a hospital laboratory for the last year, at this stage a realistic environment for drop coating deposition spectroscopy to be employed is in the laboratory rather than a GP surgery. For this technique to become automated in a laboratory a method for analysing the consistency of the blood for optimal ring formation would need to be deployed. If information regarding the cellular contents of blood could be used to identify the appropriate dilution for droplet deposition, the process of individually assessing each droplet visually prior to analysis could be eliminated. Several samples could be analysed simultaneously by depositing whole blood onto a slide using an automated multichannel micropipette. This would also improve the quality of the droplet by reducing manual errors. Spectroscopic measurements of each sample could be taken within seconds and results uploaded onto the hospital database. From a sample

being taken at a GP surgery, the sample would be sent to the laboratory the same day and analysed within hours.

There is potential for this technique to be improved further so that a hand held device could measure droplets of blood taken directly from a patient allowing the method to be introduced into the GP surgery. However this requires a lot more research, with more samples and experimentation.

7.5. Future Work

As already highlighted, larger sample groups with more spectra are required to improve the performance of the classification models. This is particularly true for the blood droplet Raman model in order to prove it can perform as well as the buffy coat IR model. Due to the small number of samples and spectra obtained it is impossible to have accounted for variations in the population that may have contributed to biochemical variances seen in the data such as gender, sex, smoking habits and diet etc. It may also be interesting to look at a secondary control group made up of patients who have a high white blood cell count due to an infection. Although changes in the concentrations of certain cellular components have been identified here for leukaemia diagnosis, for example the amount of DNA and protein, this would also be apparent in an infection as more white blood cells are produced thus increasing the number of immature white blood cells in circulation.

Further work to improve the blood droplet Raman model would be to look at the effects of hydration on the blood droplets to see if this causes any detectable biochemical changes that may be mistaken for changes relating to pathology. A further understanding of the consistency of the blood samples at the different dilutions would also be beneficial in creating a method where only one blood droplet was needed for qualitative spectroscopic analysis thus eliminating the need for a visual check of the samples.

In order to evaluate the use of the blood droplet Raman model for other leukaemia subtypes, it would be necessary to obtain samples from a wider range of pathologies. It would be particularly useful to look at chronic myeloid leukaemia (CML), acute lymphoblastic leukaemia (ALL) and acute myeloid leukaemia (AML) to identify the biochemical differences between cancers of the myeloid and lymphoid progenitors as well as between acute and chronic versions of the same cancer.

For this work to be translated into a hospital laboratory, multicentre trials of the method would be required in order to identify the sensitivity and specificity of the method in different populations using different spectrometers. The cost of the method is also very important, particularly in today's economic climate where hospitals are continually trying to cut costs. The upfront cost of a Raman spectrometer is very high, and the CaF₂ slides are very expensive compared to the simple glass slides that are used in the hospital laboratory. Thus there are many obstacles to overcome before clinical implementation is possible.

Other areas where this method could be deployed is monitoring of chemotherapy treatment in cancer patients. By identifying the biochemical changes that take place during treatment, the effectiveness of it could be identified.

Future work for lymphoma diagnosis would involve setting up ongoing sample collection with a head and neck surgeon within the hospital. The pathologies of the samples collected would need to be recorded on a password-protected database by the health professionals involved, like the leukaemia samples were. Samples preparation involving water rather than phosphate buffered saline would also have to be explored and once an appropriate protocol is set in place, IR and Raman analysis followed by multivariate statistical analysis would need to be carried out to identify biochemical markers for the diagnosis of lymphomas.

REFERENCES

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K. & Walter, P. (2002) *Molecular Biology of The Cell* (4th ed), Garland Science: New York.

Alfano, R. R., Tata, D. B., Cordero, J., Tomashefsky, P., Longo, F. W. & Alfano, M. A. (1984) Laser Induced Fluorescence Spectroscopy from Native Cancerous and Normal Tissue. *IEEE Journal of Quantum Electronics*, 20(12): 1507-1511.

Alfano, R. R., Tang, G. C., Pradhan, A., Lam, W., Choy, D. S. J. & Opher, E. (1987) Fluorescence Spectra from Malignant and Normal Human Breast and Lung Tissue. *IEEE Journal of Quantum Electronics*, 23(10): 1806-1811.

Alfano, R. R., Lui, C. H., Sha, W. L., Zhu, H. R., Akis, D. L., Cleary, J., Prudente, R. & Cellmer, E. (1991) Human Breast Tissues Studied By Infrared Fourier Transform Raman Spectroscopy. *Lasers in Life Sciences*, 4(1):23-28.

Anandasabapathy, S. (2008) Endoscopic Imaging: Emerging Optical Techniques for the Detection of Colorectal Neoplasia. *Current Opinions in Gastroenterology*, 24(1): 64-69.

Andrus, P. G. & Strickland, R. D. (1998) Cancer Grading by Fourier Transform Infrared Spectroscopy. *Biospectroscopy*, 4: 37-46.

Andrus, P. G. (2006) Cancer Monitoring by FTIR Spectroscopy. *Technology in Cancer Research and Treatment*, 5(2): 157-167.

Ansell, S. M. & Armitage, J. (2005) Non-Hodgkin's Lymphoma: Diagnosis and Treatment. *Symposium on Oncology Practise: Haematological Malignancies - Proceedings of Mayo Clinic*, 80(8): 1087-1097.

Applied Image Inc. (2013) *T-20 USAF 1951 Chart Standard Layout Product Specifications*, available at http://www.aig-imaging.com/mm5/PDF/USAF%201951%20Test%20Target%20T-20_v1-04.pdf (accessed 17th September 2013).

Atkins, P., & de Paula, J. (2010) *Atkins' Physical Chemistry* (9th ed), Oxford University Press: Oxford.

Attar, A. C. & Scadden, D. T. (2004) Regulation of Hematopoietic Stem Cell Growth. *Leukaemia*, 18: 1760-1768.

Babrah, J., McCarthy, K.P., Lush, R., Rye, A.D., Bessant, C. & Stone, N. (2007) FT-Infrared Spectroscopic Studies of Lymphoma, Lymphoid and Myeloid Leukaemias Cell Lines. *Progress in Biomedical Optics and Imaging - Proceedings of SPIE*, 6628, art. no. 66280R,.

Babrah, J., McCarthy, K., Lush, R. J., Rye, A. D., Bessant, C. & Stone, N. (2009) Fourier Transform Infrared Spectroscopic Studies of T-cell Lymphoma, B-cell Lymphoid and Myeloid Leukaemia Cell Lines. *Analyst*, 134: 763-768.

Babrah, J. (2009) *A Study of FT-IR Spectroscopy for the Identification and Classification of Haematological Malignancies* (PhD Thesis). Cranfield University, Cranfield.

Ball, D. W. (2007) The Electromagnetic Spectrum: A History. *Spectroscopy*, 21(3): 14-20.

Bassan, P., Kohler, A., Martens, H., Lee, J., Byrne, H. J., Dumas, P., Gazi, E., Brown, M., Clarke, N. & Gardner, P. (2010) Resonant Mie Scattering (RMieS) Correction of Infrared Spectra from Highly Scattering Biological Samples. *Analyst*. 135: 268-277.

Beleites, C., Steiner, G., Sowa, M. G., Baumgartner, R., Sobottka, S., Schackert, G. & Salzer, R. (2005) Classification of Human Gliomas by Infrared Imaging Spectroscopy and Chemometric Image Processing. *Vibrational Spectroscopy*, 38(1-2): 143-149.

Benedetti, E., Palatresi, M. P., Vergamini, P., Papineschi, F. & Spremolla, G. (1985) New Possibilities of Research in Chronic Lymphoblastic Leukaemia by Means of Fourier Transform Infrared Spectroscopy - II. *Leukaemia Research*, 9(8): 1001-1008.

Benedetti, E., Bramanti, E., Papineschi, F., Rossi, I. & Benedetti, E. (1997) Determination of the Relative Amounts of Nucleic Acids and Proteins in

Leukaemic and Normal Lymphocytes by Means of Fourier Transform Infrared Microspectroscopy. *Applied Spectroscopy*, 51(6): 792-797.

Benedetti, E., Bramanti, E., Papineschi, F., Vergamini, P. & Benedetti, E. (1998) An Approach to the Study of Primitive Thromocythemia (PT) Megakaryocytes by Means of Fourier Transform Infrared Microspectroscopy (FT-IR-M). *Cellular and Molecular Biology*, 44(1): 129-139.

Bennett, J. H. (1845) Case of Hypertrophy of the Spleen and Liver in which Death Took Place from Suppuration of the Blood. *Edinburgh Medical and Surgical Journal*, 64: 413-423.

Bigio, I. J. & Mourant, J. R. (1997) Ultraviolet and Visible Spectroscopies for Tissue Diagnostics: Fluorescence Spectroscopy and Elastic-Scattering Spectroscopy. *Physics in Medicine and Biology*, 42: 803-814.

Bonnier, F., Ali, S. M., Knief, P., Lambkin, H., Flynn, K., McDonagh, V., Healy, C., Lee, T. C., Lyng, F. M. & Byrne, H. J. (2012) Analysis of Human Skin Tissue by Raman Microspectroscopy: Dealing with the Background. *Vibrational Spectroscopy*, 61: 124-132.

Boydston-White, S., Gopen, T., Houser, S., Bargonetti, J. & Diem, M. (1999) Infrared Spectroscopy of Human Tissue. V. Infrared Spectroscopic Studies of Myeloid Leukaemia (ML-1) Cells at Different Phases of the Cell Cycle. *Biospectroscopy*, 5(4): 219-227.

Boydston-White, S., Chernenko, T., Regina, A., Miljkovic, M., Matthäus, C. & Diem, M. (2005) Microspectroscopy of Single Proliferating HeLa Cells. *Vibrational Spectroscopy*, 38:169-177.

Boydston-White, S., Romeo, M., Chernenko, T., Regina, A., Miljkovic, M. & Diem, M. (2006) Cell-Cycle-Dependent Variations in FTIR Micro-Spectra of Single Proliferating HeLa Cells: Principal Component and Artificial Neural Network Analysis. *Biochimica et Biophysica Acta*, 1758: 908-1914.

Brereton, R. G. (2009) *Chemometrics for Pattern Recognition*, John Wiley & Sons, Ltd: West Sussex.

Burnett, A. K. & Eden, O. B. (1997) The Treatment of Acute Leukaemia. *The Lancet*, 349: 270-275.

Burkacky, O. & Zambusch, A. (2008) "Coherent Anti-Stokes Raman Scattering (CARS) Microscopy" in: Lasch, P. & Kneipp, J. (editors), *Biomedical Vibrational Spectroscopy*, John Wiley & Sons, Inc: Hoboken, New Jersey, 209-220.

Cancer Research UK (2010a) *Leukaemia Statistics – Key Facts*, available at: <http://info.cancerresearchuk.org/cancerstats/types/leukaemia/> (accessed 28th September 2010).

Cancer Research UK (2010b) *Non-Hodgkin's Lymphoma (NHL) Statistics – Key Facts*, available at: <http://info.cancerresearchuk.org/cancerstats/types/nhl/?script=true> (accessed 28th September 2010).

Cancer Research UK (2010c) *Hodgkin's Lymphoma Statistics – UK Incidence Statistics*, available at: <http://info.cancerresearchuk.org/cancerstats/types/hodgkinslymphoma/incidence/> (accessed 28th September 2010).

Cancer Research UK (2011a) Cancer Mortality for Common Cancers, available at: <http://www.cancerresearchuk.org/cancer-info/cancerstats/mortality/cancerdeaths/> (accessed 11th February 2013).

Cancer Research UK (2011b) *Hodgkin's Lymphoma Statistics – UK Incidence Statistics*, available at: <http://info.cancerresearchuk.org/cancerstats/types/hodgkinslymphoma/incidence/> (accessed 11th February 2013).

Cancer Research UK (2012a) Cancer Incidence for all Cancers Combined <http://www.cancerresearchuk.org/cancer-info/cancerstats/incidence/all-cancers-combined/> (accessed 11th February 2013).

Cancer Research UK (2012b) Cancer Incidence for Common Cancers, available at: <http://www.cancerresearchuk.org/cancer-info/cancerstats/incidence/all-cancers-combined/> (accessed 11th February 2013).

Cancer Research UK (2013a) *Leukaemia Statistics – Key Facts*, available at: <http://info.cancerresearchuk.org/cancerstats/types/leukaemia/> (accessed 11th February 2013).

Cancer Research UK (2013b) *Non-Hodgkin's Lymphoma (NHL) Statistics – Key Facts*, available at: <http://info.cancerresearchuk.org/cancerstats/types/nhl/?script=true> (accessed 11th February 2013).

Cancer Research UK (2013c) *Types of Hodgkin Lymphoma*, available at: <http://www.cancerresearchuk.org/cancer-help/type/hodgkins-lymphoma/about/types-of-hodgkins-lymphoma> (accessed 15th October 2013).

Cancer Research UK (2013d) *Different Types of Non-Hodgkin Lymphoma*, available at: <http://www.cancerresearchuk.org/cancer-help/type/non-hodgkins-lymphoma/about/types/the-most-common-types-of-non-hodgkins-lymphoma> (accessed 15th October 2013).

Cantor, A. B. & Orkin, S. H. (2001) Haematopoietic Development: A Balancing Act. *Current Opinion in Genetics and Development*, 11(5): 513-519.

Celso, C. L., Wu, J. W., & Lin, C. P. (2009) *In Vivo* Imaging of haematopoietic Stem Cells and Their Microenvironment. *Journal of Biophotonics*, 2(11): 619-631.

Chaiken, J., Finney, W. F., Peterson, C. M., Peterson, C. P., Knudson, P. E., Weinstock, R. S. & Lein, P. (2000) Noninvasive, *In Vivo*, Tissue Modulated Near Infrared Vibrational Spectroscopic Study of Mobile and Static Tissues: Blood Chemistry. *Proceedings of SPIE*, 3918:135-143.

Chan, J. "Raman and Coherent anti-Stokes Raman Scattering (CARS) Spectroscopy and Imaging, Selected Topics in Biophotonics (EAD289)" Lawrence Livermore National Laboratory, Powerpoint Presentation (2009), available at: http://www.google.co.uk/url?sa=t&rct=j&q=&esrc=s&source=web&cd=7&ved=0CEoQFjAG&url=http%3A%2F%2Fcbst.ucdavis.edu%2Feducation%2Fcourses%2Fwinter-2009-ead-bim-289%2Flecture_pdfs%2Fchan_cars-lecturer.pdf&ei=3w9uU_9LhofQBfDHgZAP&usg=AFQjCNF7Njkm0ox-zAfMEwZoQltQ-KCckw&bvm=bv.66330100,d.d2k (accessed 27th January 2014).

- Chen, L. & Evans, J. R. G. (2010) Drying of Colloidal Droplets on Superhydrophobic Surfaces. *Journal of Colloid and Interface Science*, 351: 283-287.
- Chiriboga, L., Xie, P., Zhang, W. & Diem, M. (1997) Infrared Spectroscopy of Human Tissue. III. Spectral Differences Between Squamous and Columnar Tissue and Cells from the Human Cervix. *Biospectroscopy*, 3(4): 253-257.
- Chiriboga, L., Xie, P., Yee, H., Vigorita, V., Zarou, D., Zakim, D. & Diem, M. (1998a) Infrared Spectroscopy of Human Tissue. I. Differentiation and Maturation of Epithelial Cells in the Human Cervix. *Biospectroscopy*, 4(1): 47-53.
- Chiriboga, L., Xie, P., Vigorita, V., Zarou, D., Zakim, D. & Diem, M. (1998b) Infrared Spectroscopy of Human Tissue. II. A Comparative Study of Spectra of Biopsies of Cervical Squamous Epithelium and of Exfoliated Cervical Cells. *Biospectroscopy*, 4(1): 55-59.
- Choi, J., Gupta, S., Park, I., Lee, D. & Ye, J. C. (2007) Breast Cancer Diagnosis from Fluorescence Spectroscopy Using Support Vector Machine. *Optical tomography and Spectroscopy of Tissue VII – Proceedings of SPIE*, 6434, art. no. 64340P.
- Colley, C. S., Kazarian, S. G., Weinberg, P. D. & Lever, M. J. (2004) Spectroscopic Imaging of Arteries and Atherosclerotic Plaques. *Biopolymers*, 74(4): 328-325.
- Debatin, K. M., Stahnke, K., & Fulda, S. (2003) Apoptosis in Haematological Disorders. *Seminars in Cancer Biology*, 13: 149-158.
- Deegan, R. D., Bakajin, O., Dupont, T. F., Huber, G., Nagel, S. R. & Witten, T. A. (1997) Capillary Flow as the Cause of Ring Stains From Dried Liquid Drops. *Nature*, 389: 827-839.
- Deegan, R. D. (2000) Pattern Formation in Drying Drops. *Physical Review E*, 61 (1): 475-485.
- Deegan, R. D., Bakajin, O., Dupont, T. F., Huber, G., Nagel, S. R. & Witten, T. A. (2000) Contact Line Deposits in an Evaporating Drop. *Physical Review E*, 62 (1): 756-765.

Déléris, G. & Petibois, C. (2003) Applications of FT-IR Spectrometry to Plasma Contents Analysis and Monitoring. *Vibrational Spectroscopy*, 32: 129-136.

Diem, M., Chiriboga, L. & Yee, H. (2000) Infrared Spectroscopy of Human Cells and Tissue. VIII. Strategies for Analysis of Infrared Tissue Mapping Data and Applications to Liver Tissue. *Biopolymers (Biospectroscopy)*, 57: 282-290.

Diem, M., Chiriboga, L., Lasch, P. & Pacifico, A. (2002) IR Spectra and IR Spectral Maps of Individual Normal and Cancerous Cells. *Biopolymers (Biospectroscopy)*, 67: 349-353.

Diem, M., Romeo, M., Matthäus, C., Miljkovic, M., Miller, L. & Lasch, P. (2004) Comparison of Fourier Transform Infrared (FTIR) Spectra of Individual Cells Acquired using Synchrotron and Conventional Sources. *Infrared Physics & Technology*, 45: 331-338.

Dukor, R. K. (2002) Vibrational Spectroscopy in the Detection of Cancer, in: Chalmers, J. M. & Griffiths, P. R. (editors), *Handbook of Vibrational Spectroscopy*. Volume 5. John Wiley & Sons, Ltd.: Chichester. 3335-3361.

Ellis, D. I. & Goodacre, R. (2006) Metabolic Fingerprinting in Disease Diagnosis: Biomedical Applications of Infrared and Raman Spectroscopy. *Analyst*, 131: 875-885.

Endlicher, E., Knuechel, R., Hauser, T., Szeimies, R. M., Schölmerich, J. & Messmann, H. (2001) Endoscopic Fluorescence Detection of Low and High Grade Dysplasia in Barrett's Oesophagus Using Systemic or Local 5-Aminolevulinic Acid Sensitisation. *Gut*, 48(3): 314-319.

Enejder, A. M. K., Koo, T. W., Oh, J., Hunter, M., Sasic, S., Feld, M. S. & Horowitz, G. L. (2002) Blood Analysis by Raman Spectroscopy. *Optics Letters*, 27(22): 2004-2006.

Erukhimovitch, V., Talyshinsky, M., Souprun, Y. & Huleihel, M. (2002) Spectroscopic Characterization of Human and Mouse Primary Cells, Cell Lines and Malignant Cells. *Photochemistry and Photobiology*, 76(4): 446-451.

Erukhimovitch, V., Talyshinsky, M., Souprun, Y. & Huleihel, M. (2006) FTIR Spectroscopy Examination of Leukaemia Patients Plasma. *Vibrational Spectroscopy*, 40: 40-46.

Esmonde-White, K. A., Mandair, G. S., Esmonde-White, F. W. L., Raaii, F., Roessler, B. J. & Morris, M. (2009) Osteoarthritis Screening Using Raman Spectroscopy of Human Synovial Fluid Drops. *Optics in Bone Biology and Diagnostics – Proceedings of SPIE*, 7166, art no.71660J-1-8.

Fabian, H., Lasch, P., Boese, M. & Haensch, W. (2003) Infrared Microspectroscopic Imaging of Benign Breast Tumour Tissue Sections. *Journal of Molecular Structure*, 661-662: 411-417.

Filik, J. & Stone, N. (2007) Drop Coating Deposition Raman Spectroscopy of Protein Mixtures. *The Analyst*, 132: 544-550.

Filik, J. & Stone, N. (2008) analysis of Human Tear Fluid by Raman Spectroscopy. *Analytica Chimica Acta*, 616(2): 177-184.

Filik, J. & Stone, N. (2009) Investigation into the Protein Composition of Human Tear Fluid Using Centrifugal Filters and Drop Coating Deposition Raman Spectroscopy. *Journal of Raman Spectroscopy*, 40(2): 218-224.

Frank, C. J. & McCreery, R. L. (1995) Raman Spectroscopy of Normal and Diseased Human Breast Tissues. *Analytical Chemistry*, 67(5): 777-783.

Fredericks, P. (1995) Vibrational Spectroscopy Heads Towards Life. *Analysis*, S3-8.

Gasparri, F. & Muzio, M. (2003) Monitoring of Apoptosis of HL60 Cells by Fourier-Transform Infrared Spectroscopy. *Biochemical Journal*, 369: 239-248.

Griffiths, P. R. & de Haseth, J. A. (2007) *Fourier Transform Infrared Spectrometry*. John Wiley & Sons, Ltd: Hoboken, New Jersey.

Hehn, S. T., Grogan, T. M. & Miller, T. P. (2004) Utility of Fine-Needle Aspiration as A Diagnostic Technique in Lymphoma. *Journal of Clinical Oncology*, 22: 3046-52.

Hoffbrand, A. V., Moss, P. A. H. & Petit, J. E. (2010) *Essential Haematology* (5th ed), Blackwell Publishing Ltd: Oxford.

Horsnell, J. (2012) *The Use of Raman Spectroscopy for the Intra-Operative Assessment of Axillary Lymph Nodes in Breast Cancer* (PhD Thesis). Cranfield University, Cranfield.

Hutchings, J. (2009) *Advancing Clinical Application of Raman Spectroscopic Diagnosis of Oesophageal Premalignancies* (PhD Thesis). Cranfield University, Cranfield.

Isabelle, M., Stone, N., Barr, H., Vipond, M., Shepherd, N. & Rogers, K. (2008) Lymph Node Pathology Using Optical Spectroscopy in Cancer Diagnostics. *Spectroscopy*, 22(2-3): 97-104.

Jennings, C. D. & Foon, K. A. (1997) Recent Advances in Flow Cytometry: Application to the Diagnosis of Haematological Malignancy. *Blood*, 90(8): 2863-2892.

Jerjes, W., Swinson, B., Pickard, D., Thomas, G. J. & Hopper, C. (2004) Detection of Cervical Intranodal Metastasis in Oral Cancer Using Elastic Scattering Spectroscopy. *Oral Oncology*, 40: 673-678.

Kendall, C. A. (2002) *A Study of Raman Spectroscopy For The Early Detection and Classification of Malignancy in The Oesophageal Tissue* (PhD Thesis). Cranfield University, Cranfield.

Kendall, C., Stone N., Shephard, N., Gebos, K., Warren, B., Bennet, R. & Barr, H. (2003) Raman Spectroscopy, A Potential Tool For The Objective Identification and Classification of Neoplasia in Barret's Oesophagus. *Journal of Pathology*, 39: 118-33.

Kendall, C., Isabelle, M., Bazant-Hegemark, F., Hutchings, J., Orr, L., Babrah, J., Baker, R. & Stone, N. (2009) Vibrational Spectroscopy: A Clinical Tool For Cancer Diagnostics. *Analyst*, 134: 1029-1045.

Kočevar, K. (2005) *Physics of Surfaces and Interfaces*. Ljubljana: Jožef Stefan International Postgraduate School, available at: <http://www2.ijs.si/~kkocevar/skripta.pdf> (Accessed 24th March 2011).

Kocjan, G. (2005) BEST PRACTISE No 185 Cytological and Molecular Diagnosis of Lymphoma. *Journal of Clinical Pathology*, 58(6): 561-567.

Kopecký, V. & Baumruk, V. (2006) Structure of the Ring in Drop Coating Deposited Proteins and its Implication for Raman Spectroscopy of Biomolecules. *Vibrational Spectroscopy*, 42: 184-187.

Krafft, C. & Sergo, V. (2006) Biomedical Applications of Raman and Infrared Spectroscopy to Diagnose Tissues. *Spectroscopy*, 20(5): 195-218.

Krafft, C., Codrich, D., Pelizzo, G. & Sergo, V. (2008) Raman and FTIR Imaging of Lung Tissue: Methodology for Control Samples. *Vibrational Spectroscopy*, 46: 141-149.

Krishna, C. M., Sockalingum, G. D., Kegelaer, G., Rubin, S., Kartha, V. B. & Manfait, M. (2005) Micro-Raman Spectroscopy of Mixed Cancer Cell Populations. *Vibrational Spectroscopy*, 38(1-2): 95-100.

Krishna, C. M. *et al* (2006) Combined Fourier Transform Infrared and Raman Spectroscopic Approach for Identification of Multidrug Resistance Phenotype in Cancer Cell Lines. *Biopolymers*, 82: 462-470.

Kuby, J. (1997) *Immunology* (3rd ed), WH Freeman & Company: USA.

Kwong, Y. L. & Chan, T. K. (1988) Seminar in Haematology, Current Concepts in Haematology: 1. Normal Haemopoiesis. *Journal of the Hong Kong Medical Association*, 40(3): 171-173.

Larsson, K. & Hellgren, L. (1974) A Study of the Combined Raman and Fluorescence Scattering From Human Blood Plasma. *Experientia*, 30: 481-483.

Lasch, P., Boese, M., Pacifico, A. & Diem, M. (2002) FT-IR Spectroscopic Investigations of Single Cells on the Subcellular Level. *Vibrational Spectroscopy*, 28: 147-157.

Lasch, P. & Kneipp, J. (2008), *Biomedical Vibrational Spectroscopy*, John Wiley & Sons, Inc: Hoboken, New Jersey.

Lightfoot, T. J. & Roman, E. (2004) Causes of Childhood Leukaemia and Lymphoma. *Toxicology and Applied Pharmacology*, 199: 104-117.

Lilleyman, J. S., Hann, J. M., Stevens, R. F., Eden, O. B. & Richards, S. M. (1986) French American British (FAB) Morphological Classification of Childhood Lymphoblastic Leukaemia and its Clinical Importance. *Journal of Clinical Pathology*, 39: 998-1002.

Lin, S. Y., Li, M. J. & Cheng W. T. (2007) FT-IR and Raman Vibrational Microspectroscopies Used for Spectral Biodiagnosis of Human Tissues. *Spectroscopy*, 21(1): 1-30.

Liu, C. H., Das, B. B., Sha Glassman, W. L., Tang, G. C., Yoo, K. M., Zhu, H. R., Akins, D. L., Lubicz, S. S., Cleary, J., Prudente, R., Celmer, E., Caron, A. & Alfano, R. R. (1992) Raman, Fluorescence, and Time-Resolved Light Scattering as Optical Diagnostic Techniques to Separate Diseased and Normal Biomedical Media. *Journal of Photochemistry & Photobiology*, 16: 187-209.

Liu, K. Z., Schultz, C. P., Johnston, J. B. & Mantsch, H. H. (1997) Comparison of Infrared Spectra of CLL Cells With Their Ex-Vivo Sensitivity (MTT Assay) to Chlorambucil & Cladribine. *Leukaemia Research*, 21: 1125-1133.

Liu, K. Z., Shaw, R. A., Man, A., Dembinski, T. C. & Mantsch, H. H. (2002) Reagent Free, Simultaneous Determination of Serum Cholesterol in HDL and LDL by Infrared Spectroscopy. *Clinical Chemistry*, 48(3): 499-506.

Low-Ying, S., Shaw, R. A., Leroux, M. & Mantsch, H. H. (2002) Quantitation of Glucose and Urea in Whole Blood by Mid-Infrared Spectroscopy of Dry Films. *Vibrational Spectroscopy*, 28: 111-116.

Lyng, F. M., Faoláin, E. Ó., Conroy, J., Meade, A. D., Knief, P., Duffy, B., Hunter, M. B., Byrne, J. M., Kelehan, P. & Byrne, H. J. (2007) Vibrational Spectroscopy for Cervical Cancer Pathology, from Biochemical Analysis to Diagnostic Tool. *Experimental and Molecular Pathology*, 82(2): 121-129.

Mahadevan, A., Ramanujam, N., Mitchell, M. F. & Richards-Kortum, R. (1995) Optical Techniques For Diagnosis of Cervical Precancers; cCmparison of Raman and Fluorescence Spectroscopies. *Proceedings of SPIE*, 2388: 110-120.

Mahadevan-Jansen, A. & Richards-Kortum, R. (1996) Raman Spectroscopy For the Detection of Cancers and Pre-Cancers. *Journal of Biomedical Optics*, 1:31-70.

Mahadevan-Jansen, A., Mitchell, M. F., Ramanujam, N., Malpica, N., Thomsen, S., Utzinger, U. & Richards-Kortum, R. (1998) Near Infrared Raman Spectroscopy For *In Vitro* Detection of Cervical Precancers. *Photochemistry & Photobiology*, 68:123-132.

Manoharan, R., Wang, Y. & Feld, M. S. (1996) Histochemical Analysis of Biological Tissues Using Raman Spectroscopy. *Spectrochimica Acta Part A- Molecular & Biomolecular Spectroscopy*, 52: 215-249.

Manoharan, W., Shafer, K., Perelman, L., Wu, J., Chen, K., Deinum, G., Fitzmaurice, M., Myles, J., Crowe, J., Dasari, R. R. & Feld, M. S. (1998) Raman Spectroscopy and Fluorescence Photon Migration For Breast Cancer Diagnosis and Imaging. *Photochemistry & Photobiology*, 67(1): 15-22.

Mantsch, H. H., Choo-Smith, L. & Shaw, R. A. (2002) Vibrational Spectroscopy & Medicine: An Alliance in the Making. *Vibrational Spectroscopy*, 30: 31-41.

Martini, F. H. (2006) *Fundamentals of Anatomy & Physiology* (7th ed), Pearson Education, Inc: San Francisco, USA.

Mourant, J. R., Gibson, R. R., Johnson, T. M., Carpenter, S., Short, K. W., Yamada, Y. R. & Freya, J. P. (2003a) Methods for Measuring the Infrared Spectra of Biological Cells. *Physics in Medicine and Biology*, 48: 243-257.

Mourant, J. R., Yamada, Y. R., Carpenter, S., Dominique, L. R. & Freya, J. P. (2003b) FT-IR Spectroscopy Demonstrates Biochemical Differences in Mammalian Cell Cultures at Different Growth Stages. *Biophysical Journal*, 85(3): 1938-1947.

Mostaço-Guidolin, L. B. & Bachmann, L. (2011) Application of FTIR Spectroscopy for Identification of Blood and Leukaemia Biomarkers: A Review over the Past 15 Years. *Applied Spectroscopy Reviews*, 46: 388-404.

Nauman, D. (2008) "Perspectives of Biomedical Vibrational Spectroscopy" in: Lasch, P. & Kneipp, J. (editors), *Biomedical Vibrational Spectroscopy*, John Wiley & Sons, Inc: Hoboken, New Jersey, 1-9.

Neviliappan, S., Kan, L. F., Walter, T. T. L., Arulkumaran, S. & Wong, P. T. T. (2002) Infrared Spectral Features of Exfoliated Cervical Cells, Cervical Adenocarcinoma Tissue, and an adenocarcinoma Cell Line (SiSo). *Gynecologic Oncology*, 85: 170-174.

Office for National Statistics (2011a) *Childhood Cancer*, available at: <http://www.statistics.gov.uk/CCI/nugget.asp?ID=681&Pos=1&ColRank=2&Rank=384> (accessed 11th August 2011).

Office for National Statistics (2011b) *Cancer Survival*, available at: <http://www.statistics.gov.uk/CCI/nugget.asp?ID=861> (accessed 10th August 2011).

Office for National Statistics (2012) *Cancer Registrations in England 2010*, available at: http://www.ons.gov.uk/ons/dcp171780_263109.pdf (accessed 11th February 2013)

Orr, L. E., Christie-Brown, J., Hutchings, J. C., McCarthy, K., Rose, S., Thomas, M. & Stone, N. (2010) Raman Spectroscopy As A Tool For The Identification and Differentiation of Neoplasias Contained Within Lymph Nodes of the Head and Neck. *Head and Neck Optical Diagnostics I – Proceedings of SPIE, 7548, art no.7548C-70*.

Ortiz, C., Zhang, D., Xie, Y., Ribbe, A. E. Ben-Amotz, D. (2006) Validation of the Drop Coating Deposition Raman Method for Protein Analysis. *Analytical Biochemistry*, 353: 157-166.

Ottensmeier, C. (2001) The Classification of Lymphomas and Leukaemias. *Chemico- Biological Interactions*, 135-136: 653-664.

Otto, M. (2007) *Chemometrics: Statistics and Computer Application in analytical Chemistry* (2nd ed), Wiley-VCH: Weinheim, Germany.

- Parisse, F. & Allain, C. (1996) Shape Changes of Colloidal Suspension Droplets During Drying. *Journal de Physique II*, 6(7): 1111-1119.
- Pearce, E. I., & Tomlinson, A. (2000) Spatial Location Studies on the Chemical Composition of Human Tear Ferns. *Ophthalmic & Physiological Optics*, 20(4): 306-313.
- Petibois, C., Rigalleau, V., Melin, A., Perromat, A., Cazorla, G., Gin, H. & Déléris, G. (1999) Determination of Glucose in Dried Serum Samples by Fourier-Transform Infrared Spectroscopy. *Clinical Chemistry*, 45(9): 1530-1535.
- Petibois, C. & Déléris, G. (2006) Chemical Mapping of Tumor Progression by FT-IR Imaging: Molecular Histopathology. *TRENDS in Biotechnology*, 24(10): 455-462.
- Piller, G. J. (2001) Leukaemia – A Brief Historical Review from Ancient Times to 1950. *British Journal of Haematology*, 112: 282-292.
- Poon, K. W. C., Lyng, F. M., Knief, P., Howe, O., Meade, A. D., Curtin, J. F., Byrne, H. J. & Vaughn, J. (2012) Quantitative Reagent-Free Detection of Fibrinogen Levels in Human Blood Plasma Using Raman Spectroscopy. *Analyst*, 137: 10871814.
- Pully, V. V., Lenferink, A. T. M. & Otto, C. (2010) Time-Lapse Raman Imaging of Single Cells. *Journal of Raman Spectroscopy*, 42(2): 167-173.
- Purves, W. K., Sadava, D., Orians, G. H. & Heller, H. C. (2004) *Life: The Science of Biology* (7th ed), WH Freeman & Company: USA.
- Ramesh, J., Kapelushnik, J., Mordechai, J., Moser, A., Huleihel, M., Erukhimovitch, V., Levi, C. & Mordechai, S. (2002) Novel Methodology For The Follow-Up of Acute Lymphoblastic Leukaemia Using FTIR Microspectroscopy. *Journal of Biochemical and Biophysical Methods*, 51: 251-261.
- Ramesh, J., Huleihel, M., Mordechai, J., Moser, A., Erukhimovitch, V., Levi, C., Kapelushnik, J. & Mordechai, S. (2003) Preliminary Results of Evaluation of Progress in Chemotherapy For Childhood Leukaemia Patients Employing Fourier- Transform Infrared Microspectroscopy and Cluster Analysis. *Journal of Laboratory and Clinical Medicine*, 141(6): 385-394.

Reyes-Goddard, J. M., Barr, H. & Stone, N. (2005) Photodiagnosis Using Raman and Surface Enhanced Raman Scattering of Bodily Fluids. *Photodiagnosis & Photodynamic Therapy*, 2: 223-233.

Richards, S. J. & Jack, A. S. (2003) The Development of Integrated Haematopathology Laboratories: A New Approach to the Diagnosis of Leukaemia and Lymphoma. *Clinical and Laboratory Haematology*, 25: 337-342.

Romeo, M. J. & Diem, M. (2005) Infrared Spectral Imaging of Lymph Nodes: strategies for analysis and Artifact Reduction. *Vibrational Spectroscopy*, 38(1-2): 115-119.

Romeo, M. J., Boydston-White, S., Matthäus, C., Miljkovic, M., Bird, B., Chernenko, T., Lasch, P. & Diem, M. (2008) Infrared and Raman Microspectroscopic Studies of Individual Human Cells, in: Diem, M., Griffiths, P. R. & Chalmers, J. M. (editors), *Vibrational Spectroscopy for Medical Diagnosis*. John Wiley & Sons, Ltd: Chichester. 27-70.

Ronson, B. *et al* (2006) Locoregional Proton Radiotherapy of a Primary Cavernous Sinus Non-Hodgkin's Lymphoma: Case Report. *Technology in Cancer Research and Treatment*, 5(3): 281-284.

Rowan, S. M., Newton, M. I. & McHale, G. (1995) Evaporation of Microdroplets and The Wetting of Solid Surfaces. *The Journal of Physical Chemistry*, 99(35): 13268-13271.

Sahu, R. K. & Mordechai, S. (2005) Fourier Transform Infrared Spectroscopy in Cancer Detection. *Future Oncology*, 1(5): 635-647.

Sahu, R. K., Zelig, U., Huleihel, M., Brosh, N., Talyshinsky, M., Ben-Harosh, M., Mordechai, S. & Kapelushnik, J. (2006) Continuous Monitoring of WBC (Biochemistry) in An Adult Leukaemia Patient Using Advanced FTIR-Spectroscopy. *Leukemia Research*, 30(6): 687-693.

Santoro, A., Castagna, L. & Magagnoli, M. (2006) Lymphomas. *Update on Cancer Therapeutics*, 1: 467-473.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J. Y., White, D. J., Hartenstein, V., Eliceiri, K., Tomancak, P. & Cardona, A. (2012). Fiji: an open-source platform for biological-image analysis. *Nature Methods*, 9: 676-82.

Schultz, C. P., Liu, K. Z., Johnston, J. B. & Mantsch, H. M. (1996) Study of Chronic Lymphocytic Leukaemia Cells by FT-IR Spectroscopy and Cluster Analysis. *Leukaemia Research*, 20(8): 649-655.

Schultz, C. P., Liu, K. Z., Johnston, J. B. & Mantsch, H. M. (1997) Prognosis of Chronic Lymphocytic Leukaemia from Infrared Spectra of Lymphocytes. *Journal of Molecular Structure*, 408-409: 253-256.

Schweitzer-Stenner, R. (2006) Advances in Vibrational Spectroscopy as a Sensitive Probe of Peptide and Protein Structure: A Critical Review. *Vibrational Spectroscopy*, 42(1): 98-117.

Sefiane, K. (2010) On the Formation of Regular Patterns from Drying Droplets and their Potential Use for Bio-Medical Applications. *Journal of Bionic Engineering*, 7: S82-S93.

Segel, G. B. & Lichtman, M. A. (2004) Familial (Inherited) Leukaemia, Lymphoma, and Myeloma: An Overview. *Blood Cells, Molecules and Diseases*, 32: 246-261.

Shaw, R. A., Kotowich, S., Leroux, M. & Mantsch, H. H. (1998) Multi-Analyte Serum Analysis Using Mid-Infrared Spectroscopy. *Annals of Clinical Biochemistry*, 35: 624-632.

Shaw, R. A., Low-Ying, S., Leroux, M. & Mantsch, H. H. (2000) Toward Reagent-Free Clinical Analysis: Quantitation of Urine Urea, Creatinine and Total Protein From the Mid-Infrared Spectra of Dried Urine Films. *Clinical Chemistry*, 46: 1493-1495.

Shen, Y. C., Davies, A. G., Linfield, E. H., Elsey, T. S., Taday, P. F. & Arnone, D. D. (2003) The Use of Fourier Transform Infrared Spectroscopy for the Quantitative Determination of Glucose Concentration in Whole Blood. *Physics in Medicine and Biology*, 48: 2023-2032.

Sheng, D., Liu, X., Li, W., Wang, Y., Chen, X. & Wang, X. (2013) Distinction of Leukaemia Patients' and Healthy Persons' Serum Using FTIR Spectroscopy. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 101: 28-232.

Siebert, F. & Hildebrandt, P. (2008) *Vibrational Spectroscopy in Life Science*. Wiley-VCH Verlag GmbH & co: Weinheim, Germany.

Sigma-Aldrich (2011) Product Information Histopaque®-1077, available at: [http://www.sigmaaldrich.com/etc/medialib/docs/Sigma-Aldrich/Product Information Sheet/1/10771-research-use-pis.Par.0001.File.tmp/10771-research-use-pis.pdf](http://www.sigmaaldrich.com/etc/medialib/docs/Sigma-Aldrich/Product%20Information%20Sheet/1/10771-research-use-pis.Par.0001.File.tmp/10771-research-use-pis.pdf) (accessed 15th January 2011).

Sigma-Aldrich (2013a) Product Information Fibrinogen from Human Plasma, available at: [http://www.sigmaaldrich.com/content/dam/sigmaaldrich/docs/Sigma/Product Information Sheet/f3879pis.pdf](http://www.sigmaaldrich.com/content/dam/sigmaaldrich/docs/Sigma/Product%20Information%20Sheet/f3879pis.pdf) (accessed 5th February 2013).

Sigma-Aldrich (2013b) Product Information γ -Globulin from Human Blood, available at: [http://www.sigmaaldrich.com/content/dam/sigmaaldrich/docs/Sigma/Product Information Sheet/2/g4386pis.pdf](http://www.sigmaaldrich.com/content/dam/sigmaaldrich/docs/Sigma/Product%20Information%20Sheet/2/g4386pis.pdf) (accessed 5th February 2013).

Smith, E. & Dent, G. (2008) *Modern Raman Spectroscopy A Practical Approach*. John Wiley & Sons Ltd: Chichester, West Sussex.

Smith, A., Lightfoot, T., Simpson, J. & Roman, E. (2009) Birthweight, Sex and Childhood Cancer: A Report From the United Kingdom Childhood Cancer Study. *Cancer Epidemiology*, 33(5): 363-367.

Socrates, G. (2001) *Infrared and Raman Characteristic Group Frequencies* (3rd ed), John Wiley & Sons Ltd: Chichester, West Sussex.

Stone, N., Kendall, C., Shepherd, N., Crow, P. & Barr, H. (2002) Near-Infrared Raman Spectroscopy for the Classification of Epithelial Pre-Cancers and Cancers. *Journal of Raman Spectroscopy*, 33: 564-573.

Sule-Suso, J., Skingsley, D., Sockalingum, G. D., Kohler, A., Kegelaer, G., Manfait, M. & El Haj, A. J. (2005) FTIR Microspectroscopy as a Tool to Assess Lung Cancer Cells Response to Chemotherapy. *Vibrational Spectroscopy*, 38: 179-184.

Swinson, B., Jerjes, W., El-Maaytah, M., Norris, P. & Hopper, C. (2006) Optical Techniques in Diagnosis of Head and Neck Malignancy. *Oral Oncology*, 42: 221-228.

Szczepański, T., van der Velden, V. H. J. & van Dongen, J. J. M. (2003). Classification systems for Acute and Chronic Leukaemias. *Best Practice and Research Clinical Haematology*, 16(4): 561-582.

Taylor, J. C., Kendall, C. A., Stone, N. & Cook, T. A. (2007) Optical Adjuncts for Enhanced Colonoscopic Diagnosis. *British Journal of Surgery*, 94(1): 6-16.

Three Counties Cancer Network (2005) *Cancer Patients: A Simplified Pathway*, available at: <http://www.the3ccancernet.org.uk/downloads/Generic%20Pathway%20Master%20Apr05.pdf> (accessed 15th January 2011).

Vineis, P. *et al.* (1996) Incidence and Time Trends for Lymphomas, Leukaemias, and Myelomas: Hypothesis Generation. *Leukaemia Research*, 20(4): 285-290.

Wagnières, G. A., Star, W. M. & Wilson, B. C. (1998) *In Vivo* Fluorescence Spectroscopy and Imaging for Oncological Applications. *Photochemistry and Photobiology*, 68(5): 603-632.

Waugh, A. & Grant, A. (2005) *Ross and Wilson Anatomy and Physiology in Health and Illness* (9th ed), Elsevier Ltd: China.

Woolf, N. (1998) *Pathology Basic and Systemic*. W. B. Saunders Company Ltd: London.

Young, T. (1805) An Essay on the Cohesion of Fluids. *Philosophical Transactions of the Royal Society*, 95: 65-87.

Zelig, U., Mordechai, S., Shubinsky, G., Sahu, R. K., Huleihel, M., Leibovitz, E., Nathan, I. & Kapelushnik, J. (2011) Pre-Screening and Follow-Up of Childhood

Acute Leukaemia Using Biochemical Infrared Analysis of Peripheral Blood Mononuclear Cells. *Biochimica et Biophysica Acta* 1810: 827-835.

Zhang, D., Xie, Y., Mrozek, M. F., Ortiz, C., Davisson, V. J. & Ben-Amotz, D. (2003) Raman Detection of Proteomic Analytes. *Analytical Chemistry*, 75(21): 5703-5709.

3CCN Cancer Patients: A Simplified Pathway



1. REFERRAL: PATIENT IS REFERRED INTO HOSPITAL CARE

- Patients are most often referred by a GP. Other possibilities include referral from a screening service or through routine or emergency admission to hospital.
- Specific criteria may apply to appointments, e.g. the 'maximum 2 week wait' for urgent referrals, or the 31 and 62 day maximum targets for waiting times shown on the pathway diagram above.
- The point of entry into the pathway may be part way through the process shown, e.g. where a patient attends hospital for a routine screening, or as a result of onwards referral from a ward to other services.
- A patient may attend more than one hospital or other healthcare centre, e.g. local hospital or cancer centre.
- Where possible patient choice is provided, for example, as to which centre is attended.

Continues on next page ...

2. INITIAL SERVICES

- Patients may access services differently, for example, with regards to the order, number and types of test.
- Diagnostic services include, for example: blood tests, x-ray, biopsy, ultrasound, MRI, mammogram or other scan.

3. DIAGNOSIS & TREATMENT PLAN

- Diagnosis and Treatment Plan are discussed and agreed with each patient (the date of this consultation is the "Decision to Treat" date).
- Each patient's case is also reviewed at a team meeting of specialists to ensure best consideration of each patient's needs and the quality of care provided (i.e. a Multidisciplinary Team Meeting, or MDT).
- After a patient is given a diagnosis of cancer, the patient's GP is informed of the diagnosis by the end of the following working day.
- GP is also informed if treatment changes, and at the end of a treatment episode (e.g. upon admission to hospital, or the start of radiotherapy, or follow up care).

4. TREATMENT & SUPPORT SERVICES

- Treatment services may include surgery in theatre (or perhaps in clinic for minor procedures), chemotherapy, radiotherapy, hormonal treatment, pain relief or other palliative care, and possibly (voluntary) participation in a clinical trial.
- Patients may require repeat or additional services, e.g. from specialists in: diet and nutrition, information services, lymphoedema care, physiotherapy, psychosocial support, psychosexual counselling, speech & language therapy or symptom/pain management.

5. PATIENT PATHWAY CONTINUES

- Patient leaves NHS Hospital Trust care with provision made for appropriate follow-up appointments and/or support services, for example: GP-based care, occupational therapy, hospice, district/community nurse, palliative care, terminal care, bereavement services, or social services (e.g. home help).
- GP is informed at key points.

Supportive & Palliative Care Services: Note that these services, and others not listed here, may be accessed at any point along the patient pathway.

idmMar05

APPENDIX B

1. Characteristic Infrared Peaks in the Fingerprint region 900-1800 cm⁻¹

ν : stretching vibrations

δ : bending (scissoring) vibrations

s: symmetric

as: asymmetric

Peak	Assignment	Tissue or Substance	Reference
900-1150	ν (C-O/C-OH): Glycogen	Liver tissue	Diem 2000
900-1185	Glycogen	Cervical tissue	Sahu 2005
900-1200	Glycogen	Cervical cells	Chiriboga 1998a
900-1300	δ (C-O): saccharides, glucose, lactate, glycerol	Leukaemia plasma	Erukhimovitch 2006
900-1300	ν (C-O): saccharides, glucose, lactose, glycerol	Plasma	D��leris 2003
900-1350	ν (PO ₂ ⁻): collagen & glycogen	Malignant non-Hodgkins Lymphoma	Andrus 1998
950	DNA	Fibroblast cell lines	Mourant 2003b
950-1200	Glycogen	Cervical cells	Diem 2002
965	DNA	ALL Lymphocytes	Ramesh 2002
965	DNA	ALL Lymphocytes	Ramesh 2003
965	ν (C-C) deoxyribose: DNA & RNA	Human promyelocytic leukaemia (HL60) cell lines	Gasparri 2003
965	ν (C-O) ribose & (PO ₂ ⁻)	Nucleic acids	Chiriboga 1998a
965	(PO ₂ ⁻): DNA	Myeloid Leukaemia (ML-1) cell lines	Boydston- White 1999
966	DNA	CLL cells	Schultz 1996
966	DNA	Colonic tissue	Sahu 2005
966	Nucleic acids & lipids	White blood cells	Sahu 2006
966	DNA	CLL cells	Liu 1997
967	DNA (hydrated)	Myeloid Leukaemia (ML-1) cell lines	Boydston- White 1999
968	DNA	Bovine Liver Tissue	Lasch 2002
970	DNA	Leukaemic Lymphocytes	Benedetti 1997
970	DNA	Cervical Tissue	Nevilliappan

			2002
970	DNA	Normal human skin	Boydston-White 2005
983-1116	α_2 -macroglobulin	Sigma Aldrich	D��leris 2003
996	$\nu\delta$ ring: uracyl	Leukaemic Lymphocytes	Benedetti 1997
996	RNA	Colonic tissue	Sahu 2005
997-1062	ν (C-O): glucose	Human serum	Petibois 1999
1000	Nucleic acids (DNA:RNA)	Atheroscleratic Plaques	Colley 2004
1000-1100	ν sPO ₂ ⁻	ALL Lymphocytes	Ramesh 2002
1000-1140	Nucleic acids	Human promyelocytic leukaemia (HL60) cell lines	Gasparri 2003
1000-1150	ν sPO ₂ ⁻	ALL Lymphocytes	Ramesh 2002
1000-1150	Nucleic acids (in absence of glycogen)	Nucleic acids	Chiriboga 1998a
1000-1200	Nucleic acids (DNA:RNA)	ALL Lymphocytes	Ramesh 2002
1000-1250	DNA/RNA	Cervical cells	Diem 2002
1015	ν (C-O): ribose	Leukaemic Lymphocytes	Benedetti 1997
1015	Glycogen	Lung carcinoma (A549) cell lines	Petibois 2006
1020	s (PO ₂ ⁻): nucleic acids (DNA)	Glycogen-poor cells (colonocytes/lymphocytes) Malignant non-Hodgkins Lymphoma	Andrus 1998
1020	Nucleic acids (DNA:RNA)	ALL Lymphocytes	Ramesh 2003
1020	DNA	Colonic tissue	Sahu 2005
1020-1120	ν (C-O): sugars & complex carbohydrates	Astrocytic glioma tissue	Beleites 2005
1025	Glycogen	Normal cervix tissue	Andrus 1998
1025	CH ₂ OH & ν (C-O) coupled to δ (C-O): C-OH of carbohydrates (glucose, fructose, glycogen etc.)	Moloney murine sarcoma virus transformed mouse & human fibroblast cells, Mouse (NIH/3T3) & human (Hep-2) fibroblast cell lines, Mouse & human malignant cells	Erukhimovitch 2002
1025-1047	ν C-O or δ C-OH Glycogen	Cervical Tissue	Nevilliappan 2002
1028	Glycogen	Squamous (superficial layer) cervical epithelial	Chiriboga 1997

		tissue	
1028	$\nu(\text{C-O}/\text{C-C})$ & C-OH deformation: glycogen	Glycogen	Chiriboga 1998a
1031	Collagen	Normal epithelial tissues	Andrus 1998
1031	Nucleic acids	Squamous (basal layer) cervical epithelial tissue	Chiriboga 1998a
1033	$\nu(\text{C-O})$: D-glucose	Sigma Aldrich	Petibois 1999
1033	Glucose	Lung carcinoma (A549) cell lines	Petibois 2006
1033-1074	Cholesterol	Sigma Aldrich	Dél�ris 2003
1040	Carbohydrate moiety glycoproteins	Glandular (columnar) & squamous (superficial layer) cervical epithelial tissue	Chiriboga 1997
1050	Carbohydrate	Bovine Liver Tissue	Lasch 2002
1054	DNA (hydrated)	Myeloid Leukaemia (ML-1) cell lines	Boydston-White 1999
1056	Carbohydrate	Leukaemia plasma	Erukhimovitch 2006
1059	$\nu(\text{C-O})$: D-fructose	Sigma Aldrich	Petibois 1999
1060	(C-O) ribose & $\nu(\text{PO}_2^-)$ of DNA	Myeloid Leukaemia (ML-1) cell lines	Boydston-White 1999
1060	Nucleic acids	Squamous (basal layer) cervical epithelial tissue	Chiriboga 1998a
1063	$\nu\text{CO-O-C}$	Bovine Liver Tissue	Lasch 2002
1064	$\nu(\text{C-O})$: D-galactose	Sigma Aldrich	Petibois 1999
1064	CH_2OH & $\nu(\text{C-O})$ coupled to $\delta(\text{C-O})$: C-OH of carbohydrates (glucose, fructose, glycogen etc.)	Moloney murine sarcoma virus transformed mouse & human fibroblast cells, Mouse (NIH/3T3) & human (Hep-2) fibroblast cell lines, Mouse & human malignant cells	Erukhimovitch 2002
1065	$\nu(\text{C-O})$ ribose & (PO_2^-)	Nucleic acids	Chiriboga 1998a
1067	DNA (dried)	Myeloid Leukaemia (ML-1) cell lines	Boydston-White 1999
1070	Nucleic acids (DNA:RNA)		Mourant 2003a
1075	$\nu(\text{C-O})$: D-mannose	Sigma Aldrich	Petibois 1999
1076	Carbohydrate moiety glycoproteins	Glandular (columnar) & squamous (superficial layer) cervical epithelial	Chiriboga 1997

		tissue	
1078	(PO ₂ ⁻): RNA, DNA & phospholipids	Normal human skin	Boydston-White 2005
1078	Glycogen	Squamous (superficial layer) cervical epithelial tissue	Chiriboga 1997
1078	<i>v</i> (C-O/ C-C) & C-OH deformation: glycogen	Glycogen	Chiriboga 1998a
1078	DNA (in absence of glycogen)	Squamous (basal layer) cervical epithelial tissue	Chiriboga 1998a
1080	Nucleic acids (RNA)	Myeloid Leukaemia (ML-1) cell lines	Boydston-White 1999
1080	Nucleic acids (DNA:RNA)		Sahu 2005
1080	Nucleic acids (DNA)	Benign human skin fibroblast	Diem 2004
1080	Nucleic acids (DNA)	Cervical cells	Diem 2002
1080	Deoxyribose (C-O): DNA	Astrocytic glioma tissue	Beleites 2005
1080	Nucleic acids (DNA:RNA)	Atheroscleratic Plaques	Colley 2004
1080	<i>vs</i> PO ₂ ⁻	Lung Cancer Cells	Sule-Suso 2005
1080	<i>vs</i> PO ₂ ⁻	Cervical Tissue	Nevilliappan 2002
1080	<i>vs</i> (PO ₂ ⁻): nucleic acids	Human promyelocytic leukaemia (HL60) cell lines	Gasparri 2003
1081	Nucleic acids	Squamous (basal layer) cervical epithelial tissue	Chiriboga 1998a
1081	Carbohydrates	Lung carcinoma (A549) cell lines	Petibois 2006
1082	<i>s</i> (PO ₂ ⁻) & <i>as</i> (PO ₂ ⁻)	Moloney murine sarcoma virus transformed mouse & human fibroblast cells, Mouse (NIH/3T3) & human (Hep-2) fibroblast cell lines, Mouse & human malignant cells	Erukhimovitch 2002
1083	<i>vs</i> PO ₂ ⁻	Bovine Liver Tissue	Lasch 2002
1084	(C-O) ribose & <i>v</i> (PO ₂ ⁻) of DNA	Myeloid Leukaemia (ML-1) cell lines	Boydston-White 1999
1084	<i>s</i> (PO ₂ ⁻): phosphate/sugar backbone nucleic acids	Malignant non-Hodgkins Lymphoma	Andrus 1998

1085	(C-O) ribose & ν (PO ₂ ⁻) of RNA	Myeloid Leukaemia (ML-1) cell lines	Boydston-White 1999
1087	ν SPO ₂ ⁻	CLL cells	Schultz 1996
1087	ν SPO ₂ ⁻	CLL cells	Liu 1997
1088	ν SPO ₂ ⁻	Human Serum	Liu 2002
1090	DNA (hydrated)	Myeloid Leukaemia (ML-1) cell lines	Boydston-White 1999
1091-1122	Triglyceride	Sigma Aldrich	Dél�ris 2003
1095	ν (PO ₂ ⁻)	Nucleic acids	Chiriboga 1998a
1095	(C-O) ribose & ν (PO ₂ ⁻) of DNA	Myeloid Leukaemia (ML-1) cell lines	Boydston-White 1999
1099	ν C=O (ester)	Fibroblast cell lines	Mourant 2003b
1100	Nucleic acids (DNA:RNA)		Mourant 2003a
1115-1134	Lactate	Sigma Aldrich	Dél�ris 2003
1120	Carbohydrate moiety glycoproteins	Cervical squamous epithelial tissue (superficial layer)	Chiriboga 1997
1121	Nucleic acids (DNA:RNA)	ALL Lymphocytes	Ramesh 2003
1121	ν (PO ₂ ⁻): nucleic acids (RNA)	Glycogen-poor cells (colonocytes/ lymphocytes) Malignant non-Hodgkins Lymphoma	Andrus 1998
1121	RNA	Colonic tissue	Sahu 2005
1127	Lactic acid	Lung carcinoma (A549) cell lines	Petibois 2006
1147-1186	Cholesterol esters	Sigma Aldrich	Dél�ris 2003
1150	ν (C-C/C-O): saccharides, glucose, lactate, glycerol	Bovine Liver Tissue	Lasch 2002
1151	Glycogen	Squamous (superficial layer) cervical epithelial tissue	Chiriboga 1997
1151	ν (C-O/ C-C) & C-OH deformation: glycogen	Glycogen	Chiriboga 1998a
1155	ν C-OH carbohydrate	Cervical Tissue	Nevilliappan 2002
1170	ν (C-C/C-O): saccharides, glucose, lactate, glycerol	Atheroscleratic Plaques	Colley 2004
1170	C-OH groups of serine, threonine	Cervical Tissue	Nevilliappan 2002

	& tyrosine (cell proteins)		
1170-1310	$vsPO_2^-$	ALL Lymphocytes	Ramesh 2002
1173	$vasCO-O-C$	Bovine Liver Tissue	Lasch 2002
1200-1245	$vsPO_2^-$	ALL Lymphocytes	Ramesh 2002
1200-1330	$v(C-N) \& \delta(N-H)$: amide III	Leukaemic Lymphocytes	Benedetti 1997
1204	Collagen	Normal epithelial tissues	Andrus 1998
1206	Amide III	Squamous (stromal layer) cervical epithelial tissue	Chiriboga 1998a
1225	DNA (hydrated)	Myeloid Leukaemia (ML-1) cell lines	Boydston-White 1999
1230	$as(PO_2^-)$	Nucleic acids	Chiriboga 1998a
1230	$v(C-N) \& \delta(N-H)$: amide III	Thrombocythemic megakaryocytes	Benedetti 1998
1230	DNA	Colonic tissue	Sahu 2005
1230	(PO_2^-) : DNA	Myeloid Leukaemia (ML-1) cell lines	Boydston-White 1999
1234	Nucleic acids (DNA)	Cervical cells	Diem 2002
1235	Nucleic acids (DNA)	Benign human skin fibroblast	Diem 2004
1235	(PO_2^-) : nucleic acids (amide III)	Squamous (basal layer) cervical epithelial tissue	Chiriboga 1998a
1236-1242	Collagen & nucleic acids	Malignant non-Hodgkins Lymphoma	Andrus 1998
1237	$vasPO_2^-$	Atherosclerotic Plaques	Colley 2004
1237	$vasPO_2^-$	Bovine Liver Tissue	Lasch 2002
1238	(PO_2^-) : RNA, DNA & phospholipids	Normal human skin	Boydston-White 2005
1238	Amide III	Squamous (stromal layer) cervical epithelial tissue	Chiriboga 1998a
1240	$vasPO_2^-$	CLL cells	Schultz 1996
1240	$vasPO_2^-$	Cervical Tissue	Nevilliappan 2002
1240	$vas(PO_2^-)$	Human promyelocytic leukaemia (HL60) cell lines	Gasparri 2003
1240	$vasPO_2^-$	CLL cells	Liu 1997
1240	$vasPO_2^-$	Lung Cancer Cells	Sule-Suso 2005
1240	Deoxyribose (PO_2^-) : DNA	Astrocytic glioma tissue	Beleites 2005
1240	Collagen	Malignant non-Hodgkins Lymphoma	Andrus 1998
1240-1340	$vC-N \& \delta N-H$	Bovine Liver Tissue	Lasch 2002

	Amide III		
1242	Amide III	Bovine Serum Albumin	Chiriboga 1998a
1242	$\nu_s(\text{PO}_2^-)$	Human Serum	Liu 2002
1242	Amide III	Normal epithelial tissue	Andrus 1998
1243 (1200-1400)	$s(\text{PO}_2^-)$ & $as(\text{PO}_2^-)$	Moloney murine sarcoma virus transformed mouse & human fibroblast cells, Mouse (NIH/3T3) & human (Hep-2) fibroblast cell lines, Mouse & human malignant cells	Erukhimovitch 2002
1244	RNA	Colonic tissue	Sahu 2005
1245	Amide III	Myeloid Leukaemia (ML-1) cell lines	Boydston-White 1999
1254	Amide III	Astrocytic glioma tissue	Beleites 2005
1270	$\nu(\text{C-N})$ & $\delta(\text{N-H})$: amide III	Leukaemia plasma	Erukhimovitch 2006
1279-1337	α_1 -acid glycoprotein	Sigma Aldrich	Déléris 2003
1280	Amide III	Squamous (stromal layer) cervical epithelial tissue	Chiriboga 1998a
1282	Collagen	Normal epithelial tissues	Andrus 1998
1292-1338	Transferin	Sigma Aldrich	Déléris 2003
1300	Glycogen	Fibroblast cell lines	Mourant 2003b
1300	$\nu\text{C-N}$ & $\delta\text{N-H}$ Amide III	Lung Cancer Cells	Sule-Suso 2005
1307	Amide III	Bovine Serum Albumin	Chiriboga 1998a
1310	Protein side chains	Myeloid Leukaemia (ML-1) cell lines	Boydston-White 1999
1310	Peptide side chains	Astrocytic glioma tissue	Beleites 2005
1316-1419	Immunoglobulin-G ₁	Sigma Aldrich	Déléris 2003
1337	Collagen	Normal epithelial tissues	Andrus 1998
1352	$\delta_s(\text{CH}_3)$ $\delta_s(\text{CH}_2)$	Bovine Liver Tissue	Lasch 2002
1360-1428	Immunoglobulin-M	Sigma Aldrich	Déléris 2003
1360-1430	(COO ⁻): amino acids	Leukaemia plasma	Erukhimovitch 2006
1360-1430	$\nu(\text{COO}^-)$: amino acids	Plasma	Déléris 2003
1360-1430	Amino acids	Sigma Aldrich	Déléris 2003
1363-1428	Fibrinogen	Sigma Aldrich	Déléris 2003
1372-1418	Immunoglobulin-	Sigma Aldrich	Déléris 2003

	A		
1378	$\delta s(\text{CH}_3)$ & $\delta s(\text{CH}_2)$	Human Serum	Liu 2002
1378-1428	Haptoglobin	Sigma Aldrich	Déléris 2003
1390	Protein side chains	Myeloid Leukaemia (ML-1) cell lines	Boydston-White 1999
1395	Aliphatic side chains, amino acids	Bovine Serum Albumin	Chiriboga 1998a
1397	$\nu(\text{C-C/ C-O})$: saccharides, glucose, lactate, glycerol	CLL cells	Schultz 1996
1397	$\nu(\text{COO}^-)$: amino acids	Bovine Liver Tissue	Lasch 2002
1400	δNH_2 (amino acids)	Fibroblast cell lines	Mourant 2003b
1400	Protein	White blood cells	Sahu 2006
1400-1455	$\delta a s(\text{CH}_3)$, $\delta a s(\text{CH}_2)$: fatty acids, phospholipids & triglyceride $\delta s(\text{CH}_2)$	Cervical Tissue	Nevilliappan 2002
1427-1484	Apolipoprotein- A_1	Sigma Aldrich	Déléris 2003
1430-1480	fatty acids, phospholipids & triglyceride	Leukaemia plasma	Erukhimovitch 2006
1430-1480	$\delta a s(\text{CH}_3)$, $\delta a s(\text{CH}_2)$, $\delta s(\text{CH}_3)$, $\delta s(\text{CH}_2)$: fatty acids, phospholipids & triglyceride $\delta s(\text{CH}_2)$	Plasma	Déléris 2003
1446	$\delta a s(\text{CH}_3)$, $\delta a s(\text{CH}_2)$: fatty acids, phospholipids & triglyceride $\delta s(\text{CH}_2)$	Human Serum	Liu 2002
1448	$\delta a s(\text{CH}_3)$, $\delta a s(\text{CH}_2)$: fatty acids, phospholipids & triglyceride $\delta s(\text{CH}_2)$	Atheroscleratic Plaques	Colley 2004
1450	CH_3 deformation	Lung Cancer Cells	Sule-Suso 2005
1450	Protein side chains	Myeloid Leukaemia (ML-1) cell lines	Boydston-White 1999

1452	$\delta_{as}(\text{CH}_3)$, $\delta_{as}(\text{CH}_2)$: fatty acids, phospholipids & triglyceride $\delta_s(\text{CH}_2)$	CLL cells	Liu 1997
1455	$as\text{CH}_3$ deformation	CLL cells	Schultz 1996
1455	$as\text{CH}_3$ deformation	Bovine Liver Tissue	Lasch 2002
1455	Aliphatic side chains, amino acids	Bovine Serum Albumin	Chiriboga 1998a
1464-1479	Immunoglobulin-D	Sigma Aldrich	Déléris 2003
1467	$\delta_{as}(\text{CH}_3)$, $\delta_{as}(\text{CH}_2)$: fatty acids, phospholipids & triglyceride $\delta_s(\text{CH}_2)$	CLL cells	Liu 1997
1467	δCH_2 lipid acyl	Human Serum	Liu 2002
1468	δCH_2 lipid acyl	CLL cells	Schultz 1996
1468	$\delta_{as}(\text{CH}_3)$, $\delta_{as}(\text{CH}_2)$: fatty acids, phospholipids & triglyceride $\delta_s(\text{CH}_2)$	Bovine Liver Tissue	Lasch 2002
1468	$\delta(\text{CH}_2)$: acyl chain	Astrocytic glioma tissue	Beleites 2005
1470	Peptide side chains	Astrocytic glioma tissue	Beleites 2005
1480-1600	$\delta(\text{NH})$: amide II, α -helix, proteins	Plasma	Déléris 2003
1483-1595	Amide II	Human promyelocytic leukaemia (HL60) cell lines	Gasparri 2003
1488-1600	Albumin	Sigma Aldrich	Déléris 2003
1500-1700	Purine & pyrimidine: DNA & RNA	Lung carcinoma (A549) cell lines	Petibois 2006
1505-1538	Immunoglobulin-G ₄	Sigma Aldrich	Déléris 2003
1518	$\delta(\text{N-H})$ proteins: tyrosine ring	Human glioblastoma (U87) cell lines	Petibois 2006
1518	$\delta(\text{N-H})$ collagen: tyrosine ring	Type I collagen	Petibois 2006
1530	Coupled $\nu(\text{C-N})$ & (CNH) deformation:	Myeloid Leukaemia (ML-1) cell lines	Boydston-White 1999

	amide II, proteins		
1530	δ (NH): amide II, α -helix, proteins	Atheroscleratic Plaques	Colley 2004
1530	δ (NH): amide II	Astrocytic glioma tissue	Beleites 2005
1530-1545	δ (NH): amide II, α -helix, proteins	Leukaemic Lymphocytes	Benedetti 1997
1530-1570	δ (NH): amide II, α -helix, proteins	Bovine Liver Tissue	Lasch 2002
1538	ν (CN) & δ (CNH): amide II	Human promyelocytic leukaemia (HL60) cell lines	Gasparri 2003
1540	δ (NH): amide II, α -helix, proteins	Lung Cancer Cells	Sule-Suso 2005
1540	Amide II	Cervical Adenocarcinoma (HeLa) cell lines	Boydston-White 2005
1541	Amide II	Bovine Serum Albumin	Chiriboga 1998a
1542	δ (NH): amide II, α -helix, proteins	CLL cells	Liu 1997
1544	ν (CN) & δ (CNH): amide II	Moloney murine sarcoma virus transformed mouse & human fibroblast cells, Mouse (NIH/3T3) & human (Hep-2) fibroblast cell lines, Mouse & human malignant cells	Erukhimovitch 2002
1544	ν (C-N) & δ (CNH): amide II	Leukaemia plasma	Erukhimovitch 2006
1545	δ (NH): amide II, α -helix, proteins	Cervical Tissue	Nevilliappan 2002
1545	δ (NH): amide II, α -helix, proteins	Colonic Tissue	Sahu 2005
1546	δ (NH): amide II, α -helix, proteins	Human Serum	Liu 2002
1546	δ (N-H) proteins: amide II	Human glioblastoma (U87) cell lines	Petibois 2006
1546	δ (N-H) collagen: amide II	Type I collagen	Petibois 2006
1549	δ (NH): amide II, α -helix, proteins	Thrombocytic megakaryocytes	Benedetti 1998
1550	δ (NH): amide II	Liver tissue	Diem 2000
1550	Amide II	Cervical cells	Diem 2002
1550	Amide II	Squamous (stromal layer) cervical epithelial tissue	Chiriboga 1998a
1554	δ (N-H) collagen:	Type I collagen	Petibois 2006

	amide II		
1560-1630	δNH_2 (amino acids)	Plasma	Déléris 2003
1571	$\delta(\text{N-H})$ collagen: amide II	Type I collagen	Petibois 2006
1575	$\delta(\text{N-H})$ proteins: amide II	Human glioblastoma (U87) cell lines	Petibois 2006
1580-1700	Nucleic acid aromatic base breathing, $\nu(\text{C=O})$		
1592	δNH_2 (amino acids)	Leukaemia plasma	Erukhimovitch 2006
1592	$\delta(\text{NH}_2)$ collagen: amine	Type I collagen	Petibois 2006
1595	$\delta(\text{NH}_2)$ proteins: amine	Human glioblastoma (U87) cell lines	Petibois 2006
1599-1710 (1640)	$\nu(\text{C=O})$ coupled to $\delta(\text{N-H})$ & $\nu(\text{C-N})$: amide I	Human promyelocytic leukaemia (HL60) cell lines	Gasparri 2003
1600-1720	$\nu(\text{C=O})$: amide I, β -sheet: proteins, turns, coils	Plasma	Déléris 2003
1613-1637 (1633)	$\nu(\text{C=O})$: amide I, β -sheet	Human promyelocytic leukaemia (HL60) cell lines	Gasparri 2003
1615	$\nu(\text{C=O})$ proteins: β -turn	Human glioblastoma (U87) cell lines	Petibois 2006
1617	$\nu(\text{C=O})$ collagen: β -turn	Type I collagen	Petibois 2006
1620	$\nu(\text{C=O})$ nucleic acids bases & ring breathing modes	Nucleic acids	Chiriboga 1998a
1620-1690	$\nu(\text{C=O})$: amide I	Fibroblast cell lines	Mourant 2003b
1620-1690	$\nu(\text{C=O})$: amide I	Bovine Liver Tissue	Lasch 2002
1622-1652	Immunoglobulin-G ₂	Sigma Aldrich	Déléris 2003
1628-1652	Immunoglobulin-G ₃	Sigma Aldrich	Déléris 2003
1628	$\nu(\text{C=O})$ collagen: parallel β -strand	Type I collagen	Petibois 2006
1630	$\nu(\text{C=O})$ proteins: parallel β -strand	Human glioblastoma (U87) cell lines	Petibois 2006
1632-1652	Urea	Sigma Aldrich	Déléris 2003
1635	$\nu(\text{C=O})$: amide I	Atherosclerotic Plaques	Colley 2004
1637-1645	$\nu(\text{C=O})$: amide I, coils	Human promyelocytic leukaemia (HL60) cell lines	Gasparri 2003

1638	$\nu(\text{C}=\text{O})$ collagen: triple helix	Type I collagen	Petibois 2006
1640-1660	$\nu(\text{C}=\text{O})$: amide I	Leukaemic Lymphocytes	Benedetti 1997
1642	$\nu(\text{C}=\text{O})$ proteins: unordered structure	Human glioblastoma (U87) cell lines	Petibois 2006
1643	$\nu(\text{C}=\text{O})$: amide I	Moloney murine sarcoma virus transformed mouse & human fibroblast cells, Mouse (NIH/3T3) & human (Hep-2) fibroblast cell lines, Mouse & human malignant cells	Erukhimovitch 2002
1643	$\nu(\text{C}=\text{O})$: amide I	Leukaemia plasma	Erukhimovitch 2006
1645	Amide I	Cervical Adenocarcinoma (HeLa) cell lines	Boydston- White 2005
1645-1662 (1650)	$\nu(\text{C}=\text{O})$: amide I, α -helix	Human promyelocytic leukaemia (HL60) cell lines	Gasparri 2003
1647	$\nu(\text{C}=\text{O})$ collagen: unordered structure	Type I collagen	Petibois 2006
1650	$\nu(\text{C}=\text{O})$: amide I	Myeloid Leukaemia (ML-1) cell lines	Boydston- White 1999
1650	Amide I	Benign human skin fibroblast	Diem 2004
1650	$\nu(\text{C}=\text{O})$: amide I	Liver tissue	Diem 2000
1650	$\nu(\text{C}=\text{O})$: amide I	Astrocytic glioma tissue	Beleites 2005
1652	$\nu(\text{C}=\text{O})$: amide I	CLL cells	Liu 1997
1652-1656	$\nu(\text{C}=\text{O})$: amide I	CLL cells	Schultz 1996
1654	$\nu(\text{C}=\text{O})$ proteins: α -helix	Human glioblastoma (U87) cell lines	Petibois 2006
1655	$\nu(\text{C}=\text{O})$: amide I	Cervical Tissue	Nevilliappan 2002
1655	Amide I	Cervical Cells	Chiriboga 1998b
1655	$\nu(\text{C}=\text{O})$: amide I	Human Serum	Liu 2002
1655	$\nu(\text{C}=\text{O})$: amide I	Thrombocytic megakaryocytes	Benedetti 1998
1655	$\nu(\text{C}=\text{O})$ nucleic acids bases & ring breathing modes	Nucleic acids	Chiriboga 1998a
1666	$\nu(\text{C}=\text{O})$ collagen: α -helix	Type I collagen	Petibois 2006
1657	$\nu(\text{C}=\text{O})$: amide I		Sahu 2005
1658	Amide I	Bovine Serum Albumin	Chiriboga

			1998a
1660-1676	Apolipoprotein-C ₃	Sigma Aldrich	Déléris 2003
1662-1682	$\nu(\text{C=O})$: amide I, turns	Human promyelocytic leukaemia (HL60) cell lines	Gasparri 2003
1666	$\nu(\text{C=O})$ proteins: β -turn	Human glioblastoma (U87) cell lines	Petibois 2006
1668	$\nu(\text{C=O})$ collagen: β -turn	Type I collagen	Petibois 2006
1678	$\nu(\text{C=O})$ proteins: parallel β -strand	Human glioblastoma (U87) cell lines	Petibois 2006
1679	$\nu(\text{C=O})$ collagen: parallel β -strand	Type I collagen	Petibois 2006
1690	$\nu(\text{C=O})$ nucleic acids bases & ring breathing modes	Nucleic acids	Chiriboga 1998a
1690	$\nu(\text{C=O})$ proteins: anti-parallel β -strand	Human glioblastoma (U87) cell lines	Petibois 2006
1690	$\nu(\text{C=O})$ collagen: anti-parallel β -strand	Type I collagen	Petibois 2006
1700-1800	Cholesterol	Human Serum	Liu 2002
1705-1960	RNA $\nu(\text{CCO})$	Leukaemic Lymphocytes	Benedetti 1997
1715	$\nu(\text{C=O})$ collagen: esters	Type I collagen	Petibois 2006
1718	$\nu(\text{C=O})$ proteins: lipid esters	Human glioblastoma (U87) cell lines	Petibois 2006
1730	$\nu(\text{C=O})$ lipid ester	Astrocytic glioma tissue	Beleites 2005
1732-1739	$\nu(\text{C=O})$ lipids, cholesterol esters, triglycerides	Plasma	Déléris 2003
1735	$\nu(\text{C=O})$: lipids	Moloney murine sarcoma virus transformed mouse & human fibroblast cells, Mouse (NIH/3T3) & human (Hep-2) fibroblast cell lines, Mouse & human malignant cells	Erukhimovitch 2002
1737	$\nu(\text{C=O})$ lipids, cholesterol esters, triglyceride	Bovine Liver Tissue	Lasch 2002
1738	Phospholipids	Cervical Adenocarcinoma (HeLa) cell lines	Boydston-White 2005

1739	$\nu(\text{C}=\text{O})$ proteins: acid esters	Human glioblastoma (U87) cell lines	Petibois 2006
1740	$\nu(\text{C}=\text{O})$ lipids, cholesterol esters, triglyceride	Bovine Liver Tissue	Lasch 2002
1740	Lipid	White blood cells	Sahu 2006
1740	$\nu(\text{C}=\text{O})$ lipids	Malignant non-Hodgkins Lymphoma	Andrus 1998
1740	$\nu\text{s}(\text{PO}_4^{3-})$ & $\nu\text{as}(\text{PO}_4^{3-})$: phospholipids	Cervical cells	Diem 2002
1742	$\nu(\text{C}=\text{O})$ lipids, cholesterol esters, triglyceride	Fibroblast cell lines	Mourant 2003b

2. Characteristic Raman Shifts in the Fingerprint region 900-1800 cm⁻¹

ν : stretching vibrations

δ : bending (scissoring) vibrations

s: symmetric

as: asymmetric

Shift cm ⁻¹	Assignment	Tissue or Substance	Reference
917	Deoxyribose (CH ₂ deformation)	DNA (B)	Mahadevan-Jansen 1996
920	Glucose	Sigma	Mahadevan-jansen 1998
920	ν C-C elastin	Sigma Aldrich E1625	Kendall 2002
920	C-C proline ring	Type I collagen (Human placenta) and infiltrating ductal carcinoma (breast)	Frank 1995
920	ν C-C proline ring collagen	Sigma Aldrich C7774	Kendall 2002
920	ν C-C proline ring/glucose/lactate	Oesophagus tissue	Stone 2002
921	ν C-C proline ring	Collagen	Mahadevan-Jansen 1996
922	ν C-C	Cervical Tissue	Lyng 2007
928-940	C-C proline, valine		Mahadevan-jansen 1998
930	ν C-C	Tear fluid	Filik 2009
934	C-C proline ring	Type I collagen (human placenta)	Frank 1995
935-945	(C-C) skeletal vibrations, α -helix	α -helix protein secondary structure	Mahadevan-Jansen 1996
935	ν C-C proline, valine & protein backbone (α -helix)/ glycogen	Oesophagus tissue	Stone 2002
936	ν C-C, backbone in collagen	Sigma Aldrich C7774	Kendall 2002
937	Glycogen	Sigma Aldrich G0885	Kendall 2002
937	C-C proline ring	Infiltrating ductal carcinoma (breast)	Frank 1995
938	C-C backbone	Collagen	Mahadevan-Jansen 1996
938	ν C-C protein backbone	Human skin tissue	Bonnier 2012
939	ν C-C-N	Bovine serum albumin	Fredericks 1995
939	ν C-C backbone (α -helix conformation)	Blood Plasma	Poon 2012

940	vC-C-N bovine albumin	Sigma Aldrich A2153	Kendall 2002
958	CH ₃ deformation	Blood Plasma	Poon 2012
966	Desmosine & isodesmosine (amino acids)	Elastin	Manoharan 1996
972	C-C proline ring	Normal breast tissue, oleic acid methyl ester	Frank 1995
975	Deoxyribose (CH ₂ deformation)	DNA (B)	Mahadevan-Jansen 1996
978	vs phosphate ion	Phospholipids, glucose-1-phosphate	Mahadevan-jansen 1998
978	vs phosphate ion	Human cervix	Mahadevan-Jansen 1998b
978	Phosphorylated proteins & nucleic acids	Human cervical biopsies	Mahadevan 1995
1000	Phenylalanine in bovine albumin	Sigma Aldrich A2153	Kendall 2002
1000-1100	Glucose	Blood	Chaiken 2000
1000-1150	C-C skeletal	Phospholipid membrane, hydrophobic chains	Mahadevan-Jansen 1996
1001	Phenylalanine	Sigma Aldrich T8324	Kendall 2002
1001	s ring breathing phenylalanine	Oesophagus tissue	Stone 2002
1002	(C-C) skeletal, β -sheet	β -sheet protein secondary structure	Mahadevan-Jansen 1996
1002	Hydroxyproline, tyrosine	Type I collagen (human placenta)	Frank 1995
1002	Phenylalanine	Peripheral blood lymphocytes	Pully 2010
1002	Phenylalanine	Tear fluid	Filik 2008
1003	s ring breathing phenylalanine	Blood Plasma	Poon 2012
1004	Phenylalanine	Infiltrating ductal carcinoma (breast)	Frank 1995
1004	Phenylalanine & tryptophan aromatic ring	Bovine serum albumin	Fredericks 1995
1004	CC aromatic ring breathing	Cervical Tissue	Lyng 2007
1004	Phenylalanine in collagen	Sigma Aldrich C7774	Kendall 2002
1004	Phenylalanine	Albumin, chicken egg Sigma A5253	Kopecky 2006
1004-1006	CH ₃ plane rocking in carotenoids	Peripheral blood lymphocytes	Pully 2010
1005	β -carotene	Breast	Frank 1995

1006	Phenylalanine	Collagen, histones	Mahadevan-Jansen 1996
1006	Phenylalanine phenol ring	Tear fluid	Filik 2009
1014	Tryptophan		Mahadevan-Jansen 1996
1014	Tryptophan indole ring	Tear fluid	Filik 2009
1030	Collagen	Sigma Aldrich C7774	Kendall 2002
1030	δ CH phenylalanine	Human skin tissue	Bonnier 2012
1031	δ C-H phenylalanine	Oesophagus tissue	Stone 2002
1032	Proline	Type I collagen (human placenta)	Frank 1995
1032	δ CH phenylalanine	Blood Plasma	Poon 2012
1033	Phenylalanine	Albumin, chicken egg Sigma A5253	Kopecky 2006
1034	ν C-C	Cervical Tissue	Lyng 2007
1043	Proline	Infiltrating ductal carcinoma (breast)	Frank 1995
1048	Glycogen	Sigma Aldrich G0885	Kendall 2002
1065	ν C-N	Cervical Tissue	Lyng 2007
1066	ν C-C, Proline	Normal breast tissue	Frank 1995
1067	Proline	Oleic acid methyl ester	Frank 1995
1069	Hydroxyapatite shifted due to environment	Human cervix	Mahadevan-jansen 1998
1070	vs phosphate ion	Glucose-1-phosphate, collagen	Mahadevan-jansen 1998
1070	vs phosphate ion	Human cervix	Mahadevan-jansen 1998
1070	Collagen/ elastin		Mahadevan 1995
1078	ν C-N	Benign breast tissue	Alfano 1991
1079	ν C-C	Normal breast tissue	Frank 1995
1082		Oleic acid methyl ester	Frank 1995
1082	ν C-N	Blood Plasma	Poon 2012
1083	Glycogen	Sigma Aldrich G0885	Kendall 2002
1083	ν C-N proteins (& lipid)	Oesophagus tissue	Stone 2002
1090	PO_2^-	DNA (C)	Mahadevan-Jansen 1996
1090	vs PO_2^-	Peripheral blood lymphocytes	Pully 2010
1091	PO_2^-	DNA (B)	Mahadevan-Jansen 1996
1095	ν_s of two ionised phosphate oxygen's in diphosphate ester	DNA (B)	Mahadevan-Jansen 1996

1095	PO ₂ ⁻	DNA (Z), RNA (Z)	Mahadevan-Jansen 1996
1096	ν C-C	Cervical Tissue	Lyng 2007
1098	ν C-C	Cervical Tissue	Lyng 2007
1099	PO ₂ ⁻	DNA (A), RNA (A)	Mahadevan-Jansen 1996
1099	ν C-N	Human skin tissue	Bonnier 2012
1100	ν_s of two ionised phosphate oxygen's in diphosphate ester	DNA (A)	Mahadevan-Jansen 1996
1100-1110	(C-C) skeletal vibrations	Unordered protein secondary structure	Mahadevan-Jansen 1996
1002	ν CC	Cervical Tissue	Lyng 2007
1108	Desmosine & isodesmosine	Elastin	Manoharan 1996
1118	ν C-C	Normal breast tissue	Frank 1995
1119	ν C-C	Oleic acid methyl ester	Frank 1995
1123	Glycogen	Sigma Aldrich G0885	Kendall 2002
1123	Glucose	Sigma	Mahadevan-jansen 1998
1123	ν C-C lipids/ ν C-N protein	Oesophagus tissue	Stone 2002
1124	ν CC skeletal <i>trans</i>	Cervical Tissue	Lyng 2007
1125		Infiltrating ductal carcinoma (breast)	Frank 1995
1126	ν C-N protein	Human skin tissue	Bonnier 2012
1127		Type I collagen (human placenta)	Frank 1995
1128	ν C-N	Blood Plasma	Poon 2012
1154	C-C carotenoids	Peripheral blood lymphocytes	Pully 2010
1155	ν C-C (& C-N) protein (also carotenoid)	Oesophagus tissue	Stone 2002
1156	ν C-N	Bovine serum albumin	Fredericks 1995
1158	ν C-C, ν C-N		
1160	Carotenoid	Blood plasma	Larrson 1974
1166		Type I collagen (human placenta)	Frank 1995
1167		Infiltrating ductal carcinoma (breast)	Frank 1995
1170	δ C-H tyrosine	Oesophagus tissue	Stone 2002
1175	ν C-O	Human cervical biopsies	Mahadevan 1995
1175	δ C-H tyrosine	Blood Plasma	Poon 2012
1177	NH ₃	Human skin tissue	Bonnier 2012
1180	Tyrosine	Breast	Manoharan 1998

1205	Tyrosine & phenylalanine	Bovine albumin, Sigma Aldrich A2153	Kendall 2002
1206	Hydroxyproline, tyrosine	Type I collagen (human placenta) and infiltrating ductal carcinoma (breast)	Frank 1995
1207	Tyrosine & phenylalanine	Bovine serum albumin	Fredericks 1995
1208	Thymine	DNA (B)	Mahadevan-Jansen 1996
1208	Tryptophan & phenylalanine ν (C-C ₆ H ₅)	Blood Plasma	Poon 2012
1208	Tyrosine & Phenylalanine	Albumin, chicken egg Sigma A5253	Kopecky 2006
1209	Tryptophan & pheylalanine ν (C-C ₆ H ₅)	Oesophagus tissue	Stone 2002
1214	ν CC backbone carbon phenyl ring	Cervical Tissue	Lyng 2007
1227-1247	Amide III (ν C-N, δ N-H) β -sheet	β -sheet protein secondary structure	Mahadevan-Jansen 1996
1235-1270	Amide III (ν C-N, δ N-H) unordered		Mahadevan-Jansen 1996
1236	ν CN, δ NH amide III band	Cervical Tissue	Lyng 2007
1238	Amide III in elastin	Sigma Aldrich E1625	Kendall 2002
1239	Thymine	DNA (A), RNA (A)	Mahadevan-Jansen 1996
1240	Vibration of pyrimidine bases (cytosine and thymine)	DNA	Mahadevan-Jansen 1996
1240	Amide III - ν (C-N)	Benign breast tumour	Alfano 1991
1240	Amide III collagen	Sigma Aldrich C7774	Kendall 2002
1240	Amide III - ν (C-N)	Uterus cancer	Liu 1992
1240	ν CN, δ NH amide III band	Cervical Tissue	Lyng 2007
1240-1265	Amide III (ν C-H of proteins, mainly α -helix)	Oesophagus tissue	Stone 2002
1245-1305	Amide III(ν C-N δ N-H)	Histones	Mahadevan-Jansen 1996
1246	Amide III	Bovine serum albumin	Fredericks 1995
1246	Amide III - ν (C-N)	Collagen and DNA	Mahadevan 1995
1246	Amide III - ν (C-N)	Human cervix precancer	Mahadevan

			1995
1246	Amide III (vC-N)	Human cervical biopsies	Mahadevan 1995
1246	Amide III (vC-N)		Liu 1992
1247	Amide III	Type I collagen (human placenta) and infiltrating ductal carcinoma (breast)	Frank 1995
1247	Amide III - v(C-N)	Endometrium cancer	Liu 1992
1248	Amide III	Breast	Manoharan 1998
1248	Amide III (vC-N, δ N-H)	Collagen	Mahadevan-Jansen 1996
1250	Cytosine	DNA (B)	Mahadevan-Jansen 1996
1250	Amide III	Peripheral blood lymphocytes	Pully 2010
1254	Amide III (vC-N, δ N-H)	Elastin	Mahadevan-Jansen 1996
1254	Amide III	Tear fluid	Filik 2008
1254	Amide III	Tear fluid	Filik 2009
1256	Glycogen	Sigma Aldrich G0885	Kendall 2002
1258-1304	Amide III (vC-N, δ N-H)	α -helix protein secondary structure	Mahadevan-Jansen 1996
1260	Structural protein	Breast	Manoharan 1998
1260-1280	Amide III	Bovine serum albumin	Fredericks 1995
1262	Amide III - v(C-N)	Benign or normal cervix	Liu 1992
1262	Amide III - v(C-N)	Benign or normal uterus	Liu 1992
1262	Amide III - v(C-N)	Benign or normal endometrium	Liu 1992
1262	Amide III - v(C-N)	Benign or normal ovary	Liu 1992
1264	Amide III collagen	Sigma Aldrich C7774	Kendall 2002
1265	Amide III	Normal breast tissue	Frank 1995
1267	Amide III	Infiltrating ductal carcinoma (breast)	Frank 1995
1267	δ (=C-H)	Phospholipid membrane	Mahadevan-Jansen 1996
1268	Amide III	Oleic acid methyl ester	Frank 1995
1268	Amide III - v(C-N) and δ (NH) in bovine albumin	Sigma Aldrich A2153	Kendall 2002
1269	Amide III	Type I collagen (human placenta)	Frank 1995
1270	Amide III (vC-N)	Endometrium cancer	Liu 1992
1271	Amide III (vC-N, δ N-H)	Collagen	Mahadevan-Jansen 1996

1271	Amide III (ν C-N) α -helix	Blood Plasma	Poon 2012
1275	Tyrosine	Albumin, chicken egg Sigma A5253	Kopecky 2006
1300	ν (C-C) ν (C-N)	Benign breast tissue	Alfano 1991
1300	δ C-H	Lipid	Liu 1992
1300	Lipids – fatty acids	Breast	Manoharan 1998
1303	δ (CH ₂)	Phospholipid membrane	Mahadevan- Jansen 1996
1303	CH ₃ CH ₂ twisting	Normal breast tissue and infiltrating ductal carcinoma	Frank 1995
1304	CH ₃ CH ₂ twisting	Oleic acid methyl ester	Frank 1995
1313	CH ₃ CH ₂ twisting collagen/lipids	Oesophagus tissue	Stone 2002
1314	CH deformation	Cervical Tissue	Lyng 2007
1316	Histidine	Sigma Aldrich T8776	Kendall 2002
1316	Guanine	DNA (Z), RNA (Z)	Mahadevan- Jansen 1996
1318	Guanine	DNA (A), RNA (A)	Mahadevan- Jansen 1996
1318	C-H deformation	Blood Plasma	Poon 2012
1319	CH ₃ CH ₂ twisting	Type I collagen (human placenta)	Frank 1995
1321	In bovine albumin	Sigma Aldrich A2153	Kendall 2002
1325	Tryptophan ring vibrations		Liu 1992
1330	Tryptophan ν (C-C)	Cervix cancer	Liu 1992
1330	Tryptophan ν (C-C)	Uterus cancer	Liu 1992
1330	Tryptophan ν (C-C)	Ovary cancer	Liu 1992
1330	C-H	Nucleic acid bases and DNA, phospholipids	Mahadevan- jansen 1998
1330	C-H	Human cervix precancer	Mahadevan- jansen 1998
1333	Glycogen	Sigma Aldrich G0885	Kendall 2002
1333	Guanine	DNA (B)	Mahadevan- Jansen 1996
1334-1335	CH ₃ CH ₂ wagging nucleic acids/proteins	Peripheral blood lymphocytes	Pully 2010
1335	Adenine	DNA (A), RNA (A)	Mahadevan- Jansen 1996
1335	CH ₂ deformation	Cervical Tissue	Lyng 2007
1335	CH ₃ CH ₂ wagging of collagen & polynucleotide chain (DNA-purine bases)	Oesophagus tissue	Stone 2002

1337	CH ₂ deformation	Cervical Tissue	Lyng 2007
1338	Tryptophan		Mahadevan-Jansen 1996
1339	Adenine	DNA (B)	Mahadevan-Jansen 1996
1339	C-H deformation	Blood Plasma	Poon 2012
1343	CH ₃ CH ₂ wagging	Type I collagen (human placenta) and infiltrating ductal carcinoma (breast)	Frank 1995
1350	Glucose	Sigma	Mahadevan-jansen 1998
1361	Tryptophan		Mahadevan-Jansen 1996
1366	δ CH ₂	Cervical Tissue	Lyng 2007
1375	Tyrosine/Adenine/Guanine (ring breathing modes DNA)	Peripheral blood lymphocytes	Pully 2010
1377	Glycogen	Sigma Aldrich G0885	Kendall 2002
1401	$s\delta$ CH ₃ proteins		Mahadevan 1995
1421	Adenine/Guanine DNA/RNA	Peripheral blood lymphocytes	Pully 2010
1434	CH ₂ in lipids	Peripheral blood lymphocytes	Pully 2010
1439	CH ₃ CH ₂ deformation	Normal breast tissue	Frank 1995
1440	δ CH ₂	Cervical Tissue	Lyng 2007
1442	CH ₃ CH ₂ deformation	Oleic acid methyl ester	Frank 1995
1442	Lipids- fatty acids	Normal breast tissue	Manoharan 1998
1443	Elastin	Sigma Aldrich E1625	Kendall 2002
1445	δ (CH ₂) or δ (CH ₃)	Cervix cancer	Liu 1992
1445	δ (CH ₂) or δ (CH ₃)	Uterus cancer	Liu 1992
1445	δ (CH ₂) or δ (CH ₃)	Benign or normal cervix	Liu 1992
1445	δ (CH ₂) or δ (CH ₃)	Benign or normal endometrium	Liu 1992
1445	δ (CH ₂) or δ (CH ₃)	Benign or normal ovary	Liu 1992
1445	δ (CH ₂) or δ (CH ₃)	Human breast tissue Benign and malignant tumours	Alfano 1991
1445	δ CH ₂	Normal breast tissue	Manoharan 1998
1445	CH ₃ CH ₂ deformation in collagen	Sigma Aldrich C7774	Kendall 2002
1445-1447	δ CH ₂ proteins/lipids	Peripheral blood lymphocytes	Pully 2010
1446	δ CH ₂ proteins/ lipids	Oesophagus tissue	Stone 2002

1447	$\delta\text{CH}_3\text{CH}_2$ bovine albumin	Sigma Aldrich A2153	Kendall 2002
1447	C-H deformation	Blood Plasma	Poon 2012
1448	Deoxyribose	DNA (B)	Mahadevan-Jansen 1996
1449	δCH_2	Albumin, chicken egg Sigma A5253	Kopecky 2006
1449	$\delta\text{CH}_3\text{CH}_2$	Bovine serum albumin	Fredericks 1995
1449	CH_2CH_3 deformation	Tear fluid	Filik 2008
1449	CH_2CH_3 deformation	Tear fluid	Filik 2009
1450	Glucose	Blood	Chaiken 2000
1450	δCH_2	Diseased breast tissue	Manoharan 1998
1450	$\delta(\text{CH}_2)$ or $\delta(\text{CH}_3)$	Endometrium cancer	Liu 1992
1450		Elastin and collagen	Liu 1992
1450	CH_3CH_2 deformation	Infiltrating ductal carcinoma	Frank 1995
1451	CH_3CH_2 deformation	Type I collagen (human placenta)	Frank 1995
1451-1454	δCH_2	Histones	Mahadevan-jansen 1998
1452	Structural protein	Breast	Manoharan 1998
1453	$\delta(\text{CH}_2)$ or $\delta(\text{CH}_3)$	Ovary cancer	Liu 1992
1453	$\delta(\text{CH}_2)$ or $\delta(\text{CH}_3)$	Benign or normal uterus	Liu 1992
1454	$\delta(\text{CH}_2)$ or $\delta(\text{CH}_3)$	Elastin, collagen and phospholipids	Mahadevan-jansen 1998
1454	$\delta(\text{CH}_2)$ or $\delta(\text{CH}_3)$	Human cervix	Mahadevan-jansen 1998
1454	δCH_3 proteins		Mahadevan 1995
1455	Glycogen	Sigma Aldrich G0885	Kendall 2002
1456	δCH_3 elastin	Sigma Aldrich E1625	Kendall 2002
1460	Pentose sugar vibration due to δCH_2	DNA	Mahadevan-Jansen 1996
1462	Deoxyribose	DNA (B)	Mahadevan-Jansen 1996
1480	Vibration of purine bases (adenine and guanine)	DNA	Mahadevan-Jansen 1996
1484	CH_2 deformation	Cervical Tissue	Lyng 2007
1487-1490	Guanine (DNA)	Peripheral blood lymphocytes	Pully 2010
1505	Tyrosine	Albumin, chicken egg Sigma A5253	Kopecky 2006
1509	Phenylalanine	Bovine serum albumin	Fredericks

			1995
1520	Carotenoid	Blood plasma	Larrson 1974
1524	C=C vibration carotenoids	Peripheral blood lymphocytes	Pully 2010
1548	NH deformation; ν CN amide II band	Cervical Tissue	Lyng 2007
1548	Tryptophan	Oesophagus tissue	Stone 2002
1552	Tryptophan	Blood Plasma	Poon 2012
1553	Tryptophan		Mahadevan- Jansen 1996
1553	Tryptophan indole ring	Tear fluid	Filik 2009
1556	Tryptophan	Bovine serum albumin	Fredericks 1995
1570	Vibration of purine bases (adenine and guanine)	DNA	Mahadevan- Jansen 1996
1576	Guanine	Peripheral blood lymphocytes	Pully 2010
1578	ν C=C olefinic	Cervical Tissue	Lyng 2007
1579	Pyrimidine ring (nucleic acids) & haem protein	Oesophagus tissue	Stone 2002
1585	ν C=C	Cervical Tissue	Lyng 2007
1585	ν C=C	Blood Plasma	Poon 2012
1602	ν CO	Cervical Tissue	Lyng 2007
1603	δ C=C phenylalanine/ tyrosine	Oesophagus tissue	Stone 2002
1605	Phenylalanine phenol ring	Tear fluid	Filik 2009
1606	Phenylalanine	Bovine serum albumin	Fredericks 1995
1606	ν C=C phenylalanine & tryptophan	Blood Plasma	Poon 2012
1601	Tyrosine	Albumin, chicken egg Sigma A5253	Kopecky 2006
1607	Phenylalanine in bovine albumin	Sigma Aldrich A2153	Kendall 2002
1616	Tyrosine	Bovine serum albumin	Fredericks 1995
1616	ν C=C of tyrosine/ tryptophan	Oesophagus tissue	Stone 2002
1616	ν C=C tyrosine & tryptophan	Blood Plasma	Poon 2012
1642	Amide I ν (C=O)	Collagen	Mahadevan- Jansen 1996
1648-1661	Amide I ν (C=O)	Histones	Mahadevan- Jansen 1996

1650-1666	C=C vibration cis - isomer in fatty acid chain	Phospholipid membrane	Mahadevan-Jansen 1996
1651	Amide I (ν C=O)	Benign breast tissue and malignant tumours	Alfano 1991
1651	Amide I (ν C=O) in collagen	Sigma Aldrich C7774	Kendall 2002
1651	Amide I (ν C=O)	Ovarian cancer	Liu 1992
1653	Amide I (ν C=O) in bovine albumin	Sigma Aldrich A2153	Kendall 2002
1654	Amide I	Normal breast tissue	Frank 1995
1654-1662	Amide I (ν C=O) α -helix	α -helix protein secondary structure	Mahadevan-Jansen 1996
1654-1685	Amide I (ν C=O)	Unordered protein secondary structure	Mahadevan-Jansen 1996
1655	Amide I	Oleic acid methyl ester	Frank 1995
1655	Amide I (ν C=O)	Endometrial cancer	Liu 1992
1655	Amide I (ν C=O proteins, α -helix conformation)/ ν C=C lipid	Oesophagus tissue	Stone 2002
1656	Amide I (ν C=O)	Collagen	Mahadevan-Jansen 1998
1656	Amide I (ν C=O)	Human cervix	Mahadevan-Jansen 1996
1656	Amide I (ν C=O)		Mahadevan 1995
1656	Amide I (ν C=O)	Benign or normal endometrium	Liu 1992
1657	Amide I (ν C=O)	Cervical cancer	Liu 1992
1657	Lipids - fatty acids	Breast	Manoharan 1998
1657	Amide I	Infiltrating ductal carcinoma (breast)	Frank 1995
1657	Fatty Acids	Peripheral blood lymphocytes	Pully 2010
1659	Amide I (ν C=O)	Benign breast tumours	Alfano 1991
1659	Amide I (ν C=O)	Benign or normal ovary, cervix, uterus	Liu 1992
1659	Amide I (ν C=O)	Uterus cancer	Liu 1992
1659	Amide I (ν C=O)	Bovine serum albumin	Fredericks 1995
1659	Amide I (ν C=O) α -helix	Blood Plasma	Poon 2012
1659	Amide I	Tear fluid	Filik 2008
1659	Amide I	Tear fluid	Filik 2009
1660-1665	Amide I (ν C=O) α -helix	Cervical Tissue	Lyng 2007

1661	Amide I ($\nu\text{C=O}$)in elastin	Sigma Aldrich E1625	Kendall 2002
1661	Amide I	Peripheral blood lymphocytes	Pully 2010
1662	Nucleic acids	Colon	Manoharan 1996
1665	Amide I	Type I collagen (human placenta)	Frank 1995
1665	Amide I	Albumin, chicken egg Sigma A5253	Kopecky 2006
1665-1680	Amide I ($\nu\text{C=O}$)	β -sheet protein secondary structure	Mahadevan-Jansen 1996
1667	Structural protein	Malignant breast	Manoharan 1998
1668	Amide I ($\nu\text{C=O}$)	Elastin	Mahadevan-Jansen 1996
1670	Amide I ($\nu\text{C=O}$)	Pure protein - β -pleated sheet	Mahadevan-Jansen 1996
1670	Amide I ($\nu\text{C=O}$)	Collagen	Mahadevan-Jansen 1996
1670-1680	C=C vibration trans-isomer in fatty acid chain	Phospholipid membrane	Mahadevan-Jansen 1996
1674	Amide I	Peripheral blood lymphocytes	Pully 2010
1683	Amide I ($\nu\text{C=O}$)	Bound protein - regular characterized secondary structure	Mahadevan-Jansen 1996
1743		Oleic acid methyl ester, normal human breast tissue	Frank 1995

APPENDIX C

Correspondence from South West REC re ethical approval:

South West 5 REC

formerly Frenchay REC
C/o North Bristol NHS Trust
Beaufort House
Southmead Hospital
Westbury-on-Trym
Bristol
BS10 5NB

Tel: 0117 323 5211

Fax: 0117 323 2832

29 April 2010

Prof Nicholas Stone
Biophotonics Research Unit
Gloucestershire Royal Hospital
Great Western Rd
Gloucester
GL1 3NN

Dear Prof Stone

Study title:	FT-Infrared Spectroscopy and Raman Spectroscopy for the Enhanced Diagnosis of Leukaemias and Lymphomas
REC reference:	06/Q2005/120
Amendment number:	3
Amendment date:	25 March 2010

The above amendment was reviewed [by the Sub-Committee in correspondence](#).

Ethical opinion

The members of the Committee taking part in the review gave a favourable ethical opinion of the amendment on the basis described in the notice of amendment form and supporting documentation.

Approved documents

The documents reviewed and approved at the meeting were:

Document	Version	Date
Notice of Substantial Amendment (non-CTIMPs)	3	25 March 2010
Covering Letter		25 March 2010

Membership of the Committee

The members of the Committee who took part in the review are listed on the attached sheet.

R&D approval

All investigators and research collaborators in the NHS should notify the R&D office for the relevant NHS care organisation of this amendment and check whether it affects R&D approval of the research.

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees (July 2001) and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

**06/Q2005/120:
correspondence**

Please quote this number on all

Yours sincerely

**Mr Anthony Sack
Committee Co-ordinator**

E-mail: Anthony.Sack@nbt.nhs.uk

<i>Enclosures:</i>	<i>List of names and professions of members who took part in the review</i>
<i>Copy to:</i>	<i>Mark Walker</i>

South West 5 REC

Attendance at Sub-Committee of the REC meeting on 14 April 2010

Also in attendance:

<i>Name</i>	<i>Position (or reason for attending)</i>
Mrs Kate McMahon-Parkes	Senior Lecturer in Adult Nursing
Mr. Anthony Sack	Frenchay REC Co-ordinator
Dr Mike Shere	Associate Specialist Breast Clinician Committee Cha

SIGMA-ALDRICH

sigma-aldrich.com

3050 Center Street, St. Louis, MO 63103 USA
Tel: (800) 551-8858 (RM) 77-4755 Fax: (800) 525-8057 (RM) 771-4755
e-mail: techserv@sigmaaldrich.com info@sigmaaldrich.com

Product Information

Histopaque®-1077

Catalog Number **10771**

Storage Temperature 2–8 °C

Product Description

Histopaque®-1077 is a sterile, endotoxin tested solution of polysucrose and sodium diatrizoate, adjusted to a density of 1.077 g/mL. This ready-to-use medium facilitates rapid recovery of viable lymphocytes and other mononuclear cells from small volumes of whole blood. Histopaque-1077 is suitable for human lymphocyte antigen (HLA) typing¹ and as the initial isolation step prior to enumeration of T, B, and 'null' lymphocytes.² It may also be employed in the preparation of pure lymphocyte suspensions for cell culture and cytotoxicity assays.³

Histopaque-1077 is a sterile solution of polysucrose, 57 g/L, and sodium diatrizoate, 90 g/L.

Density: 1.076–1.078 g/mL

Endotoxin: ≤0.3 EU/mL

pH: 8.8–9.0

Reagents and Equipment Required but Not Provided

- Centrifuge (swinging bucket rotor) capable of generating 400 × g
- Centrifuge tubes, 15 mL plastic, conical
- Isotonic phosphate buffered saline solution or appropriate cell culture medium

Precautions and Disclaimer

This product is for R&D use only, not for drug, household, or other uses. Please consult the Material Safety Data Sheet for information regarding hazards and safe handling practices.

Preparation Instructions

Specimen Collection - Collect blood in preservative-free anticoagulant (EDTA or heparin) or use defibrinated blood. For best results, blood should be processed within 2 hours.

Storage/Stability

Store the product at 2–8 °C protected from light. Histopaque-1077 has an expiration period of 3 years. Reagent label bears expiration date.

Procedure

Anticoagulated blood is layered onto Histopaque-1077. During centrifugation, erythrocytes are aggregated by polysucrose and rapidly sediment. Granulocytes become slightly hypertonic, which increases their sedimentation rate, resulting in pelleting at the bottom of the centrifuge tube. Lymphocytes and other mononuclear cells remain at the plasma/Histopaque-1077 interface. Erythrocyte contamination is negligible. Most extraneous platelets are removed by low speed centrifugation during the washing steps.

1. To a 15-mL conical centrifuge tube, add 3 mL of Histopaque-1077 and bring to room temperature.
2. Carefully layer 3 mL of whole blood onto the Histopaque-1077.
3. Centrifuge at 400 × g for exactly 30 minutes at room temperature. Centrifugation at lower temperatures, such as 4 °C, may result in cell clumping and poor recovery.
4. After centrifugation, carefully aspirate the upper layer with a Pasteur pipette to within 0.5 cm of the opaque interface containing mononuclear cells. Discard upper layer.
5. Carefully transfer the opaque interface with a Pasteur pipette into a clean conical centrifuge tube.
6. Wash the cells by adding 10 mL of isotonic phosphate buffered saline solution or appropriate cell culture medium, and mix by gently drawing in and out of a Pasteur pipette.
7. Centrifuge at 250 × g for 10 minutes.
8. Aspirate the supernatant and discard.
9. Resuspend cell pellet with 5 mL of isotonic phosphate buffered saline solution or appropriate cell culture medium, and mix by gently drawing in and out of a Pasteur pipette.

10. Centrifuge at $250 \times g$ for 10 minutes.
11. Repeat steps 8, 9 and 10, discard supernatant and resuspend cell pellet in 0.5 mL of isotonic phosphate buffered saline solution or appropriate cell culture medium.

Erythrocytes and granulocytes should pellet to the bottom of the centrifuge tube. Mononuclear cells should band at the interface between the Histopaque-1077 and the plasma. If observed results vary from expected results, please contact Sigma-Aldrich Technical Service for assistance.

References

1. Amos, D.B., and Pool, P., "HLA typing" in Manual of Clinical Immunology, Rose, N.R., and Friedman, H., eds., American Society for Microbiology, (Washington, DC: 1976) pp. 797-804.
2. Winchester, R.J., and Ross, G., "Methods for enumerating lymphocyte populations" in Manual of Clinical Immunology, Rose, N.R., and Friedman, H. eds., American Society for Microbiology, (Washington, DC: 1976) pp. 64-76.
3. Thorsby, E., and Bratlie, A., "A rapid method for preparation of pure lymphocyte suspensions." Histocompatibility Testing, Terasaki, P.I., ed., 665-666 (1970).

Histopaque is a registered trademark of Sigma-Aldrich® Biotechnology LP and Sigma-Aldrich Co.

MF,MAM 02/11-1

Sigma brand products are sold through Sigma-Aldrich, Inc.
Sigma-Aldrich, Inc. warrants that its products conform to the information contained in this and other Sigma-Aldrich publications. Purchaser must determine the suitability of the product(s) for their particular use. Additional terms and conditions may apply. Please see reverse side of the invoice or packing slip.

APPENDIX E

Pure Protein Spectra

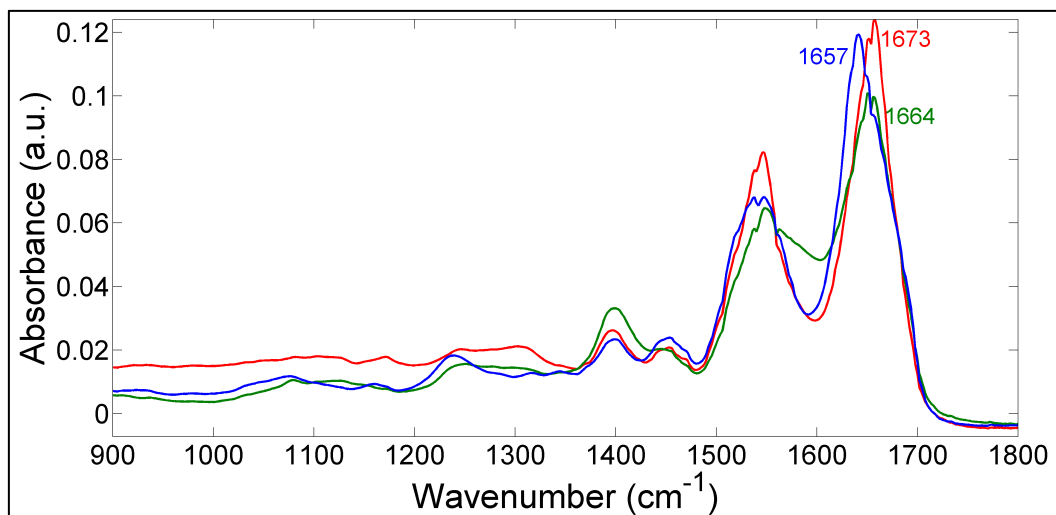


Figure 1. IR Transmittance Pure Protein Spectra

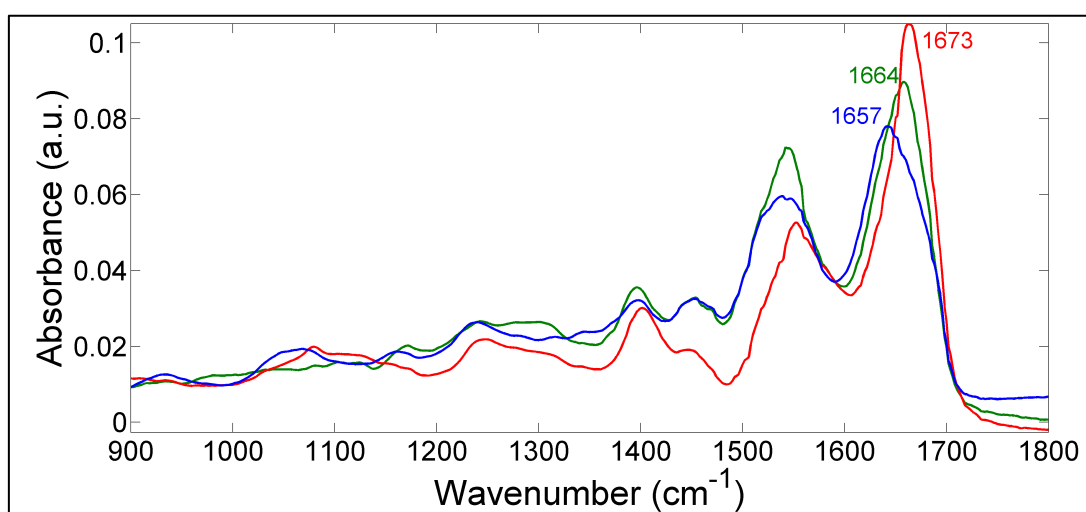


Figure 2. IR Reflectance Pure Protein Spectra

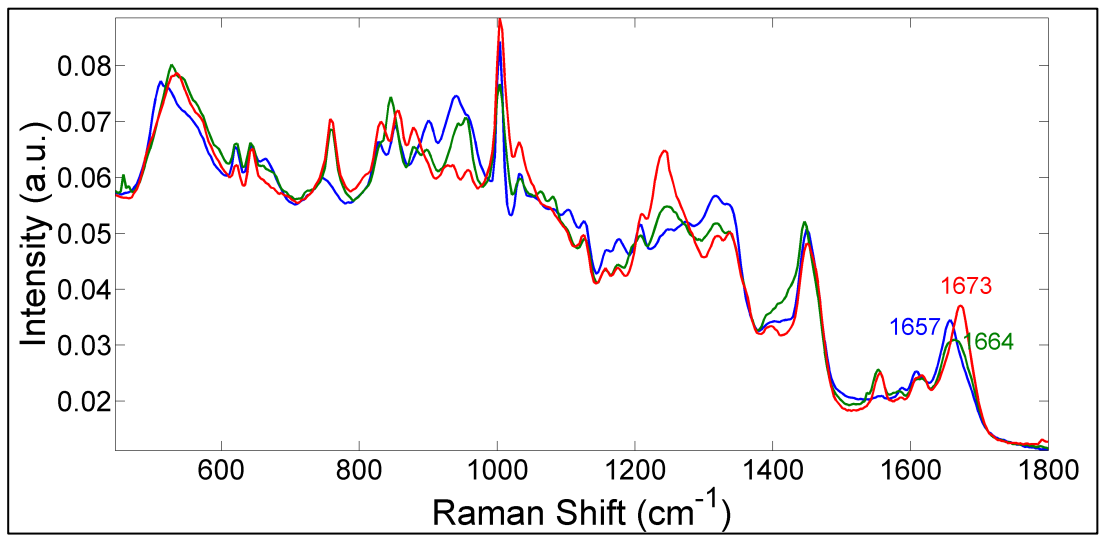


Figure 3. Raman Pure Protein Spectra

APPENDIX F

Mean Spectra

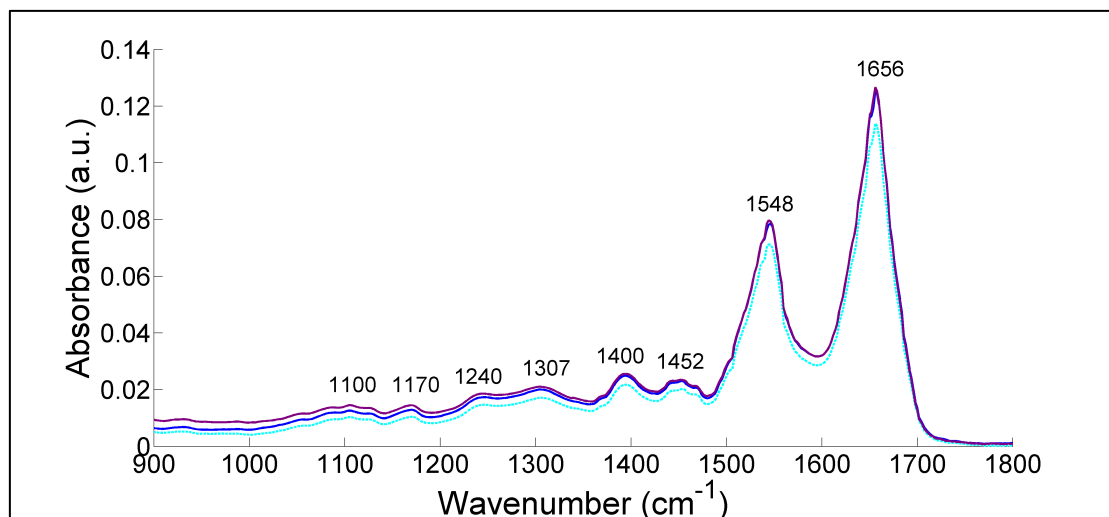


Figure 1. Mean Blood Cytospin IR Spectra of the Three Pathologies; purple: healthy, dark blue: CLL (PUP), light blue: CLL (PTP)

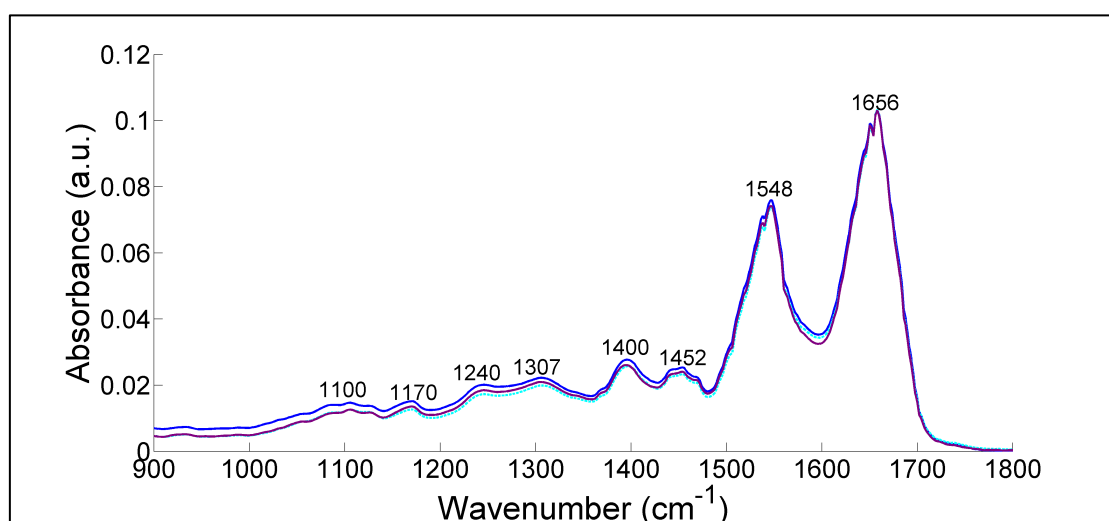


Figure 2. Mean Blood Droplet IR Spectra of the Three Pathologies; purple: healthy, dark blue: CLL (PUP), light blue: CLL (PTP)

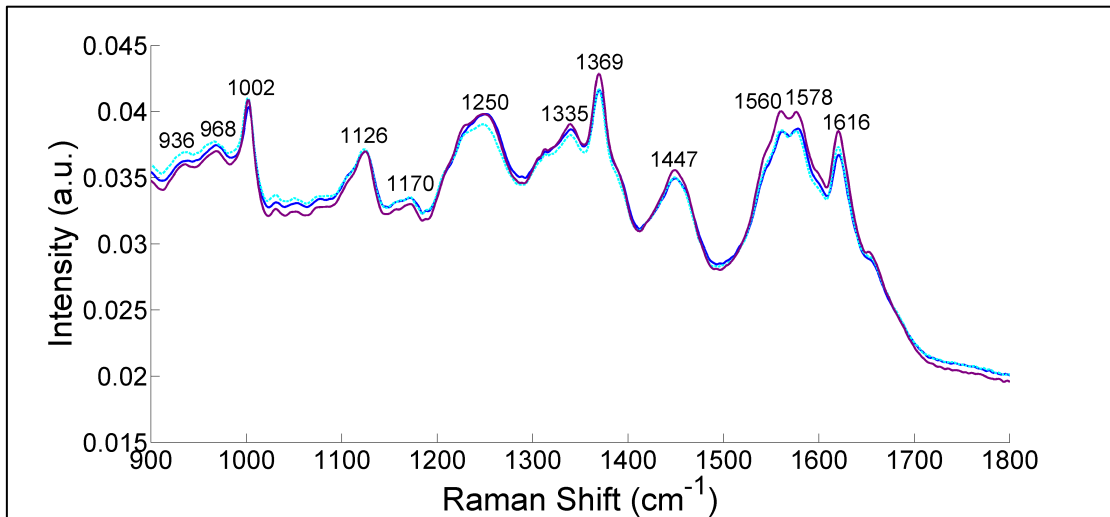


Figure 3. Mean Blood Cytospin Raman Spectra of the Three Pathologies;
purple: healthy, dark blue: CLL (PUP), light blue: CLL (PTP)

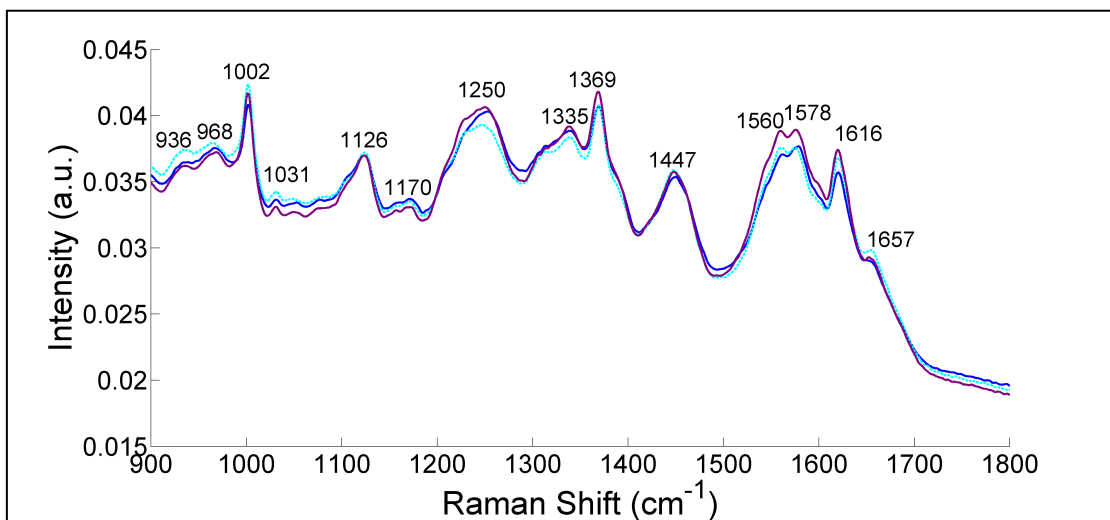


Figure 4. Mean Blood Droplet Raman Spectra of the Three Pathologies;
purple: healthy, dark blue: CLL (PUP), light blue: CLL (PTP)

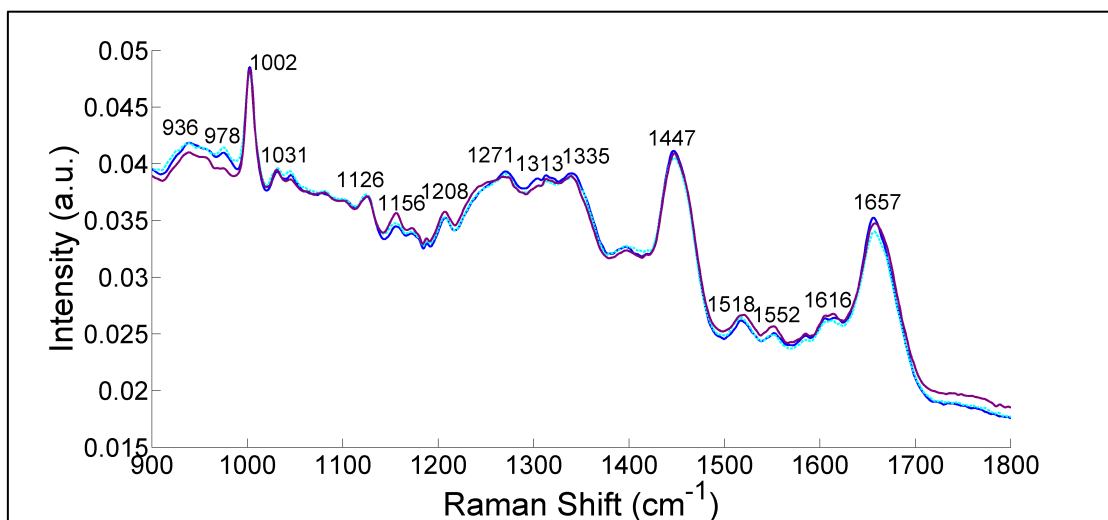


Figure 5. Mean Plasma Raman Spectra of the Three Pathologies; purple: healthy, dark blue: CLL (PUP), light blue: CLL (PTP)

APPENDIX G

List of Publications, Presentations and Prizes

Publications

Jackson, O, Lloyd, G. L. & Stone, N. (2014) Drop Coating Deposition Spectroscopy of the Blood Protein Albumin, *Journal of Vibrational Spectroscopy* (in edit).

Jackson, O., Rye, A., Lush, R., McCarthy, K., Bessant, C. & Stone, N. (2014) Vibrational Spectroscopy of Blood Components for Leukaemia Diagnosis (in edit).

Presentations

SPEC 2010

Vibrational Imaging of Whole Blood (poster presentation)

BIGSS 2010

Vibrational Spectroscopy for the Rapid and Early Diagnosis of Acute and Chronic Leukaemias and Lymphomas (poster presentation)

Medical School Open Day at GHNHSFT 2010

Early Diagnosis of Leukaemia and Lymphoma Using Raman and FTIR Spectroscopy (oral presentation)

Cranfield/ Gloucester Translational Medicine Day 2011

Raman and Fourier Transform Infrared Spectroscopy for the Diagnosis of Leukaemia and Lymphoma (oral presentation)

LINC meeting of trustees 2011

Raman and Fourier Transform Infrared Spectroscopy for the Diagnosis of Leukaemia and Lymphoma (oral presentation)

Cranfield Post-Graduate Conference 2011

Vibrational Spectroscopy for the Rapid and Early Diagnosis of Acute and Chronic Leukaemias and Lymphomas (poster presentation)

Haematology Department Meeting 2012

Raman and Fourier Transform Infrared Spectroscopy for the Rapid and Early Diagnosis of Leukaemia and Lymphoma (oral presentation)

Cranfield/ Gloucester Translational Medicine Day 2012

Vibrational Spectroscopies for the Advanced Diagnosis of Acute and Chronic Leukaemias (oral presentation)

BMLA 2012

Fourier Transform Infrared and Infrared Spectroscopy for Real-Time Diagnosis of Leukaemia (oral presentation)

ICORS 2012

Drop Coating Deposition Spectroscopy (DCDRS) for Leukaemia Diagnosis (poster presentation)

Cranfield Post-Graduate Conference 2012

Vibrational Spectroscopy for the Rapid and Early Diagnosis of Acute and Chronic Leukaemias and Lymphomas (oral presentation)

SPEC 2012

Drop Coating Deposition Spectroscopy (DCDS) for the Rapid Identification of Chronic Lymphoblastic Leukaemia (CLL) (oral presentation by Nick Stone)

Prizes

BMLA Travel Award 2012, registration fee waiver + £300 bursary

Laser Europe 2012 Educational Award of £1000