A NON-DESTRUCTIVE APPROACH FOR BREAST CANCER DIAGNOSIS AND PATHOLOGICAL STRATEGY USING INFRARED AND RAMAN SPECTROSCOPY

KOMAN KWOK-MAN TAM

A THESIS SUBMITTED FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

SCHOOL OF CHEMISTRY

THE UNIVERSITY OF SYDNEY

2006

The work submitted in this thesis has not been previously submitted for a degree or diploma at any other tertiary educational institution. To the best of my knowledge and belief, the information in this thesis contains no material previously published or written by any other person except where due reference is made.

Signed [redacted]

Koman Kwok-Man Tam

Dated 12 March 2007.

This thesis is dedicated to my uncle

譚仲品
Tam, Chung Peng

who lost his battle to a short-notice cancer in June 2006

ABSTRACT

Breast cancer is the most commonly diagnosed cancer among women in developed countries. The function of the breast is lactation during pregnancy and breast-feeding. The ducts and the lobules are the major sites where breast cancer develops. The current standard screening test for breast cancer is the triple assessment, which is a combination of a medical examination, an X-ray mammogram and a fine-needle aspiration biopsy. However, mammograms lack sensitivity and cannot detect cancer in women under the age of 50 due to constant changes in breast physiology during the menstrual cycle. In addition, pathology is subjective and often leads to many false positive and negative results. Alternative screening techniques, such as computed tomography and magnetic resonance spectroscopy and imaging, are becoming more frequently used.

Infrared (IR) and Raman spectroscopies allow investigation of samples at a molecular level and provide diagnostic information, which can be used to differentiate breast diseases. The primary goal in this research was to develop a common methodology that could be routinely used in histopathology and vibrational spectroscopy. This thesis is divided into three parts: optimisation, differentiation and characterisation. One hundred breast samples (43 benign, 7 *in-situ* carcinomas, 47 malignant and 3 pathologically unclassified) were obtained from routine needle or surgical biopsies with ethics approvals. The samples were stored in PBS/D_2O, analysed spectroscopically either as 'bulk' biopsies or as thin sections (cryosections or formalin-fixed paraffin-embedded (FFPE) sections), and examined by a pathologist.

To optimise the sampling techniques for IR and Raman spectroscopies, tissue fixation, processing, and sectioning techniques commonly used in the standard pathology laboratories were investigated. Cryosections were shown to be the preferred sampling techniques for immediate spectroscopic analysis, whereas for archival analysis, formalin-fixed then cryosectioned samples were the preferred sampling option. However, archived samples are normally preserved in formalin and mounted in paraffin blocks. Both IR and Raman spectra of the deparaffinised FFPE sections demonstrated that the standard deparaffinisation protocol was inadequate.

A characteristic of cancer is cellular proliferation, which may invade the surrounding breast tissues or metastasise. Spectroscopically, the proliferation results in an elevated intensities of IR and Raman bands in the 1700–1000 cm^{-1} (fingerprint) region, which are attributable to proteins, nucleic acids, and to a lesser extent, lipids. By comparing these bands, benign, *in-situ* carcinomas and malignant sections were distinguished using single-point IR microspectroscopy. Principal component analysis (PCA) and hierarchical cluster analysis (HCA), distinguished the IR spectra of various breast conditions with overall sensitivity and specificity of 90 and 88 per cent, respectively. Samples without prior knowledge of disease states were classified as benign fibrocystic changes, which were correlated to the histopathology results.

Characterisation of specific breast diseases such as ductal carcinoma *in-situ* (DCIS) and lobular carcinoma *in-situ* (LCIS) are important to medical diagnosis since both conditions have a high chance of developing into invasive breast cancer. By combining infrared imaging with multivariate statistical analysis, the major cellular components of a low-grade cribriform DCIS were identified. Raman imaging was used to characterise a small LCIS lesion using 514-nm laser excitation. The resulting images identified the nuclei and the nucleoli of the tumour cells within a lobule. Micro-calcification is another important indicator of cancer, which is usually difficult to see on a mammogram. A Type II micro-calcification (calcium hydroxyapatite) was identified in the Raman spectrum of an invasive ductal carcinoma lesion. Images based on at 1044 and 961 cm^{-1} bands illustrated that the calcification was denser in the centre and gradually decreases in concentration in the outer region.

A preliminary study using synchrotron-radiation based infrared (SR-FTIR) spectroscopy mapped a benign fibrocystic change, DCIS and IDC sections. The resulting maps displayed the distribution of diagnostic IR bands ascribable to proteins, lipids and nucleic acids, and were highly correlated to the areas of particular interest in each lesion.

Keywords: *breast cancer, histopathology, infrared spectroscopy, imaging, synchrotron radiation, Raman spectroscopy, sample optimisation, micro-calcification, principal component analysis, k-means, fuzzy c-means, hierarchical cluster analysis*

ACKNOWLEDGEMENTS

I'd like to express my gratitude to three groups of colleagues from: (1) School of Chemistry including my supervisors Prof. Peter Lay, Assoc. Prof. Robert Armstrong, and Dr. Elizabeth Carter; (2) The Institute for Magnetic Resonance Research including Dr. Carolyn Mountford, Prof. Peter Russell, Dr. Susan Dowd and Dr. Roger Bourne; and (3) Department of Histopathology including Dr. Jane Radford and Barbara Hernandez. Without their support in the past few years, I could not have any confidence on how to fix, section, stain the tissue sections, turn the laser on and point, map, image, and run any experiments! This research was supported by a grant from The University of Sydney, Cancer Research Fund and by the Australian Synchrotron Research Program, which is funded by the Commonwealth of Australia under the Major National Research Facilities Program.

A big thank you to all especially Liz, guiding me through different coffee breaks I could escape from the VSF. There are also a list of people I met throughout the years including Liz's officemate - Dr. Craig Marshall and their beloved 'daughter' Dr. Candace Coyle from The University of Texas, San Antonio. One more of the vibrational 'laser' family, Dr. Anne Rich. My research group members including Dr. Tony Bonin, Dr. Hugh Harris and Irma, they are simply the best to complement my anxiety going through the process of writing!

Family, unfortunately, is my major supporting casts and sponsors in my 10th academic year in The University of Sydney since my BSc (Hons) back to last century. As they have limited English literacy, so I have to express the following paragraph in Chinese.

我要多謝爸爸和媽媽辛辛苦苦給我這個希望完成今天的博士學位，多謝。另外,我哥哥和 Wynnie；還有我奶媽多年來的支持和鼓勵，沒有她美味的飯菜，那就沒有今天的我，謝謝你們。

Finally, I'd like to thank Leo Funnell and his family members for their unlimited support in the past years all the way from Bathurst to Port Stephens. Leo, thank you for all your spiritual and emotional support to share with me.

[redacted]

PUBLICATIONS TO DATE

Journal Publications (Manuscript in Progress)

K. K. M. Tam, P. A. Lay, R. S. Armstrong, E. A. Carter, C. E. Mountford, S. Dowd, P. Russell, H. Harris, C. I. Chen, Y. C. Lee, *Characterisation of Breast Ductal Carcinoma In-Situ using FTIR Imaging and Synchrotron FTIR Mapping.*

K. K. M. Tam, P. A. Lay, R. S. Armstrong, E. A. Carter, C. E. Mountford, P. Russell, *An Innovative Approach to the Diagnosis of Ductal Carcinoma In-Situ using a Combination of Infrared Imaging and Multivariate Statistical Analysis.*

CONFERENCE PROCEEDINGS

Australian Conference on Vibrational Spectroscopy

2003, Monash University, Oral Presentation:

K. Tam, P. Lay, R. Armstrong, E. Carter, C. Mountford, P. Russell, *Optimisation of Raman Imaging on a Lobular Carcinoma In-Situ Breast Lesion.* In 5th ACOVS, CSIRO: Monash University, 2003.

2005, The University of Sydney, Oral Presentation:

K. K. M. Tam, E. A. Carter, P. A. Lay, H. Harris, R. S. Armstrong, C. E. Mountford, S. Dowd, P. Russell, Y. C. Lee, C. I. Chen, *Characterisation of Ductal Carcinoma In-Situ Breast Cancer using IR Microspectroscopy.* In 6th ACOVS, ACOVS6 Organising Committee: The University of Sydney, 2005.

2005, The University of Sydney, Poster Presentation:

E. A. Carter, A. M. Rich, L. Wu, K. K. M. Tam, H. Harris, A. Levina, I. Mulyani, P. Lay, A. Stamfl, Y. C. Lee, C. I. Chen, *SR-FTIR Spectroscopy of Biological Sample at Beamline 14A NSRRC.* In 6th ACOVS, ACOVS6 Organising Committee: The University of Sydney, 2005.

International Conference on Raman Spectroscopy

2004, Gold Coast, Poster Presentation:

K. K. M. Tam, E. A. Carter, P. A. Lay, R. S. Armstrong, C. E. Mountford, S. Dowd, P. Russell, *Raman Spectroscopy For Breast Cancer: A Sample Processing Study.* In 19th ICORS, Gold Coast, 2004.

TABLE OF CONTENTS

PREFACE

1. BREAST CANCER

2. METHODS AND MATERIALS

3. HISTOPATHOLOGY OF BREAST TISSUES

4. SAMPLE OPTIMISATION FOR INFRARED SPECTROSCOPY

5. DIFFERENTIATION AND CHARACTERISATION OF BREAST TISSUES USING INFRARED SPECTROSCOPY

6. OPTIMISATION AND CHARACTERISATION OF BREAST TISSUES USING RAMAN SPECTROSCOPY

7. CONCLUSIONS AND FURTHER WORK

APPENDIX

LIST OF FIGURES

1. BREAST CANCER

2. MATERIALS AND METHODS

3. HISTOPATHOLOGY OF BREAST TISSUES

4. SAMPLE OPTIMISATION FOR INFRARED SPECTROSCOPY

5. CHARACTERISATION AND DIFFERENTIATION OF BREAST TISSUES USING INFRARED SPECTROSCOPY

6. OPTIMISATION AND CHARACTERISATION OF BREAST TISSUES USING RAMAN SPECTROSCOPY

7. CONCLUSIONS AND FURTHER WORK

APPENDIX

LIST OF TABLES

1. Breast Cancer

2. Materials and Methods

3. Histopathology of Breast Tissues

5. Differentiation and Characterisation of Breast Tissues using Infrared Spectroscopy

6. Optimisation and Characterisation of Breast Tissues using Raman Spectroscopy

LIST OF ABBREVIATIONS

Ag/SnO_2	Silver/Tin Oxide
AIHW	Australian Institute of Health and Welfare
ANN	Artifical neural network
ATR	Attenuated total reflection
BCS	Breast conserving surgery
BRCA	Breast cancer gene
CA	Central area
CaF_2	Calcium fluoride
CCD	Charge-coupled device
Cho	Choline
CPU	Computer processing unit
Cr	creatine
CT	Computed tomography
D_2O	Deuterated (heavy) water
DCIS	Ductal carcinoma *in-situ*
DLaTGS	Deuterated L-alanine triglycerine sulphate
DNA	Deoxyribose nucleic acid
EtOH	Ethanol
FCM	Fuzzy c-means
FFPE	Formalin-fixed paraffin-embedded
FNA	Fine-needle aspirated
FOV	Field of view
FPA	Focal-plane array
FT	Fourier-transform
FTIR	Fourier-transform Infrared
HCA	Hierarchical cluster analysis
H&E	Haematoxylin and eosin
HRT	Hormone replacement therapy
IDC	Invasive ductal carcinoma
ILC	Invasive lobular carcinoma
IMRR	Institute for Magnetic Resonance Research
IR	infrared
IRE	Internal reflection element
IRS	Internal reflection spectroscopy
KBr	Potassium bromide
KMC	k-means cluster
LCIS	Lobular carcinoma *in-situ*
LDA	Linear discrimination analysis
Liq. N_2	Liquid nitrogen
LO/LI	Lower-outer/lower-inner
MCT	Mercury-cadmium-telluride
MIR	Mid-infrared
MRI	Magnetic resonance imaging
MRS	Magnetic resonance spectroscopy
MSA	Multivariate statistical analysis
NA	Numerical aperture
Nd:YAG	Neodymium: Yttrium Aluminium Garnet
NIR	Near-infrared

OCT	Optimal cutting temperature
PBS	Phosphate buffer saline
PC	Principal components
PCA	Principal component analysis
PR	Professor Peter Russell
RNA	Ribose nucleic acid
RPAH	Royal Prince Alfred Hospital
RR	Resonance Raman
SBW	Strongly hydrogen-bonded water
SR-FTIR	Synchrotron-radiation source FTIR
S/N	Singal-to-noise ratio
TDLU	Terminal ductal-lobular unit
TNM	Tumour, Lymph Nodes and Metastasis
UO/UI	Upper-outer/upper-inner
UV	ultraviolet
v/v	Volume by volume
vib	Vibrational levels
w/w	Weight by weight
WBW	Weakly hydrogen-bonded water
ZnSe	Zinc selenide

Symbols used in band assignments

λ	Wavelength
ν	Stretching mode
δ	Deformation mode
ρ	Rocking mode
τ	Twisting mode
as	Anti-symmetric
s	Symmetric
ip	In-plane
op	Out-of-plane
cm^{-1}	Wavenumber unit

Amino acids

Phe	Phenylalanine
Pro	Proline

1. BREAST CANCER

1.1 BREAST FACTS: ANATOMY, INCIDENCE AND RISK FACTORS

The female breast (Figure 1.1) is a large complex network containing milk-producing sacs (lobules) and ducts for milk transportation [1]. This large network is surrounded by adipose (fat) and stromal (connective) tissues together with nerves, blood and lymphatic vessels [1]. The function of the breast is lactation, which involves production and expression of milk during pregnancy and breast-feeding. Throughout puberty and menopause, the appearance and composition of the breast change. Such changes vary among individuals, ethnicity and age groups.

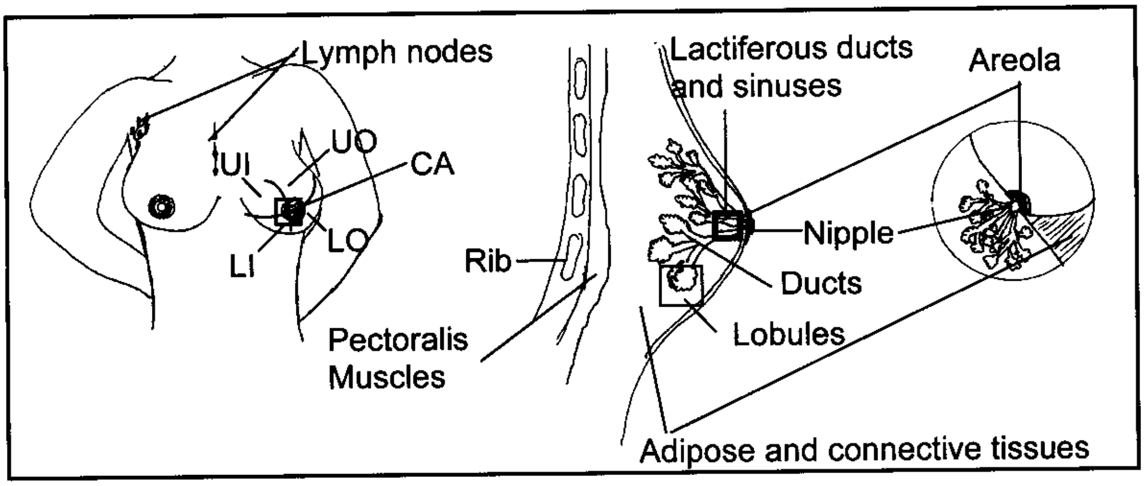


Figure 1.1. Human female breast structures and an illustration of the four quadrants UO: upper-outer; UI: upper-inner; LO: lower-outer and LI: lower-inner and the central area (CA).

1.1.1 BREAST ANATOMY

Milk is expressed in the spherical-shaped sacs called lobules. Each lobule contains a collection of small glandular spaces called acini (Figure 1.2). The milk is carried from the lobules via terminal duct-lobular units (TDLU), to lactiferous ducts and sinuses, and is subsequently emptied through the nipple during breast-feeding. The lobules and the ducts are composed of myoepithelial and cuboidal epithelial cells; these cells are surrounded by basement membranes that demarcate the loose intralobular and dense interlobular stroma (connective tissues) [1]. There are between 15 and 25 ductal systems in each breast, each with its own opening at the nipple. The nipple is surrounded by the areola, which contains modified sebaceous glands that secrete fluid to lubricate the nipple during breast-feeding [1].

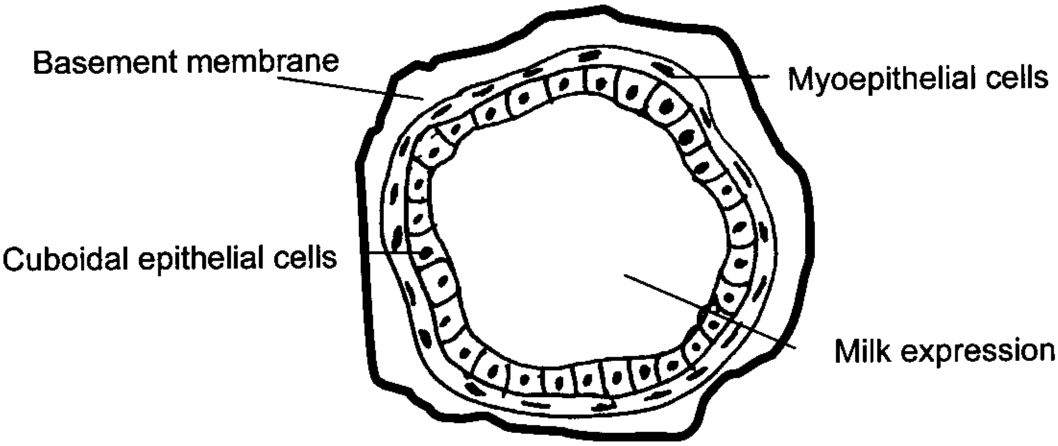


Figure 1.2. Illustration of an acinus, a milk production unit in lobules.

The breast tissues, which reach from the collarbones to the lower ribs, are interspersed by stroma (intralobular and interlobular) and adipose tissues, which comprise approximately one-third of the breast content. The breast tissues rest against the pectoral muscles and are surrounded by blood vessels, lymph nodes and vessels. Both blood and lymph vessels are involved in the metastasis of breast cancer [1-3].

Each breast can be divided into four quadrants and a central area (Figure 1.1), each division contains various proportions of breast tissue components, lymph and blood vessels. In a healthy breast, the p53 gene programs the process of apoptosis, which regulates the cell cycle. When this process malfunctions, cell growth is out of control and the accumulating tissue mass forms a tumour. As the tumour grows, it may promote blood vessel formation to supply additional oxygen and nutrients [1,4-7]. Most tumour cells originate in the ducts (80%) and, to a lesser extent, the lobules (20%) [6]. When tumour cells are found in other parts of the body, having travelled via the blood and/or lymph vessels, the process is referred to as secondary cancer or metastasis [1,4-7].

The majority of ducts and lymph nodes are located within the upper-outer (UO) quadrant and this is where approximately 50% of breast cancers are found. The central area (CA), including the nipple and areola, represents approximately 20% of cancer incidence. Both the lower-inner (LI) and lower-outer (LO) quadrants contain the least amount of breast components that are subject to tumour formation and, hence, constitute a lower incidence rate (~6% each) [6,7].

1.1.2 BREAST CANCER INCIDENCE: AUSTRALIA AND WORLDWIDE

Breast cancer is the leading female registered cancer, in both incidence and mortality, in western countries. According to The International Agency for Research on Cancer (an affiliation of the World Health Organisation), more than 1.2 million women were expected to be diagnosed with breast cancer worldwide in 2005 [8]. In Australia, 11,791 new cases were diagnosed and 2,594 deaths occurred in 2001 (last publicised figure), which accounted for 29% of cancer diagnoses and 16.3% of deaths for all of the female registered cancers [9,10]. These statistics suggest that 1 in 11 women will develop breast cancer in their lifetime with the average age of diagnosis being 59 years [9,11].

The incidence rate of breast cancer increases slightly every year in Australia (Figure 1.3), however, the overall mortality rate is steadily decreasing (2.2% p.a. between 1991 and 2001) due to improved medical screening and treatments [9-11], which highlights the need for continual improvements in diagnostics. Men are also at risk of developing breast cancer, but it accounts for < 1% (95 men) of all registered breast cancer (both sexes) in Australia. However, there is a higher mortality rate for men compared to women due to a lack of awareness of the disease [9-11].

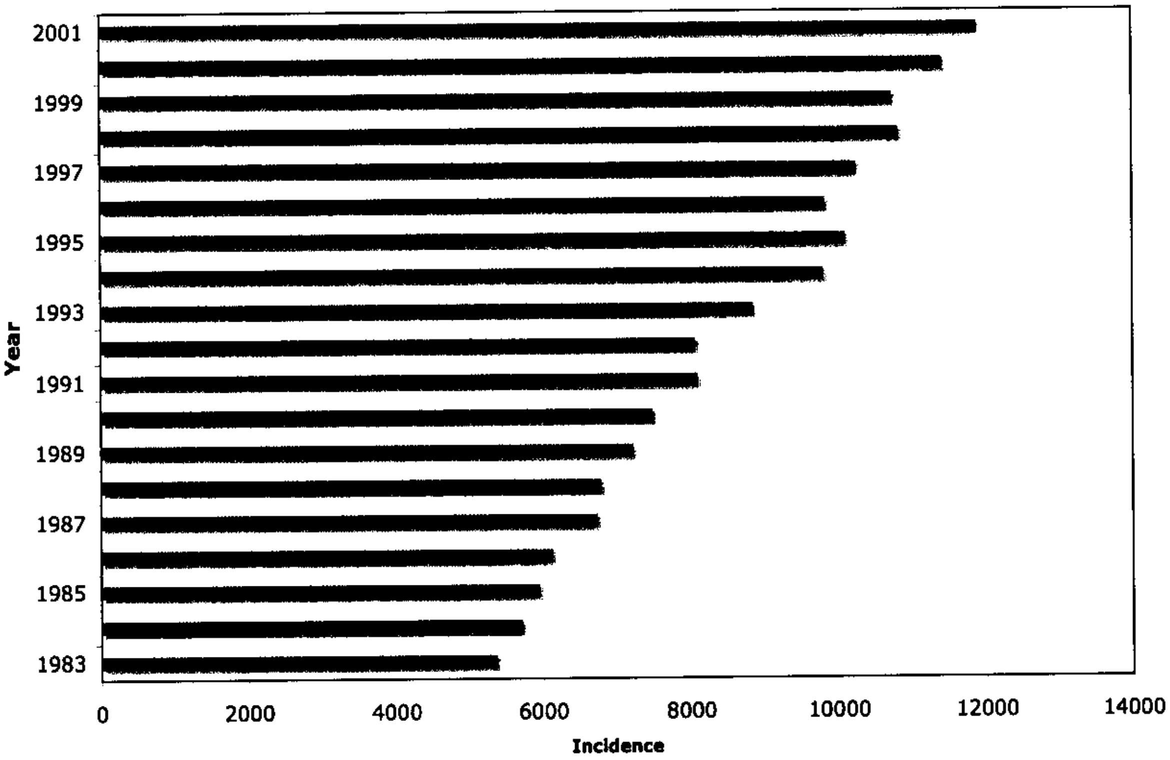


Figure 1.3. Breast cancer incidence rate in Australia from 1983 to 2001 [11].

1.1.3 BREAST CANCER: RISK FACTORS

The cause(s) of breast cancer are not fully known, however, there are a number of risk factors that might increase the likelihood of developing breast cancer. These can be categorised as controllable and uncontrollable factors and having one or more risk factors may increase the chance of, but does not necessarily lead to, cancer [1,4-7].

The uncontrollable factors that are linked to an increase in the risk of developing breast cancer are being female, age and a family history of breast cancer [1,6]. Women previously diagnosed with breast cancer, ductal carcinoma *in-situ* (DCIS) or mutations in the breast cancer genes (BRCA1 and BRCA2) are also be more vulnerable [6].

Controllable factors that may increase a woman's risk of developing breast cancer if not addressed include: the prolonged use of hormone replacement therapy (HRT), nulliparity, and bearing their first child after 30 years of age [1,6,12]. Obesity, smoking and excessive alcohol consumption are also linked to the development of breast cancer. Breast feeding, however, may lower the risk as milk is produced and expressed [12].

1.2 BREAST CANCER SCREENING: TRIPLE ASSESSMENT

Biochemical changes are at the core of all cancers and consequently identifying these changes are central to all diagnostic tests. Breast disease is first noticed as a painless lump in the breast. Other important symptoms include changes in size and shape of the breast, lump formation or thickening, dimpling of the skin and fluid or blood discharge from the nipple [6,12]. Early detection of these abnormalities increases the chance of survival.

Current screening of breast cancer is performed by a triple assessment (Figure 1.4), a combination of a routine medical examination, an X-ray mammogram and fine-needle aspiration (FNA) cytology. Clinical examination is performed to determine the symptomatic presence of dimpling and nipple deformity, enlargement of auxillary lymph nodes and loss of pliability [6]. Mammography is a specific type of imaging that uses a low-dose X-ray system for examination of breasts. Mammograms are taken by compressing the breasts between the two shields of the mammography unit and X-ray images are taken in two different views, one cranio-caudal view and one medio-lateral view [1,3]. The radiologist looks for any dense mass located within the breast.

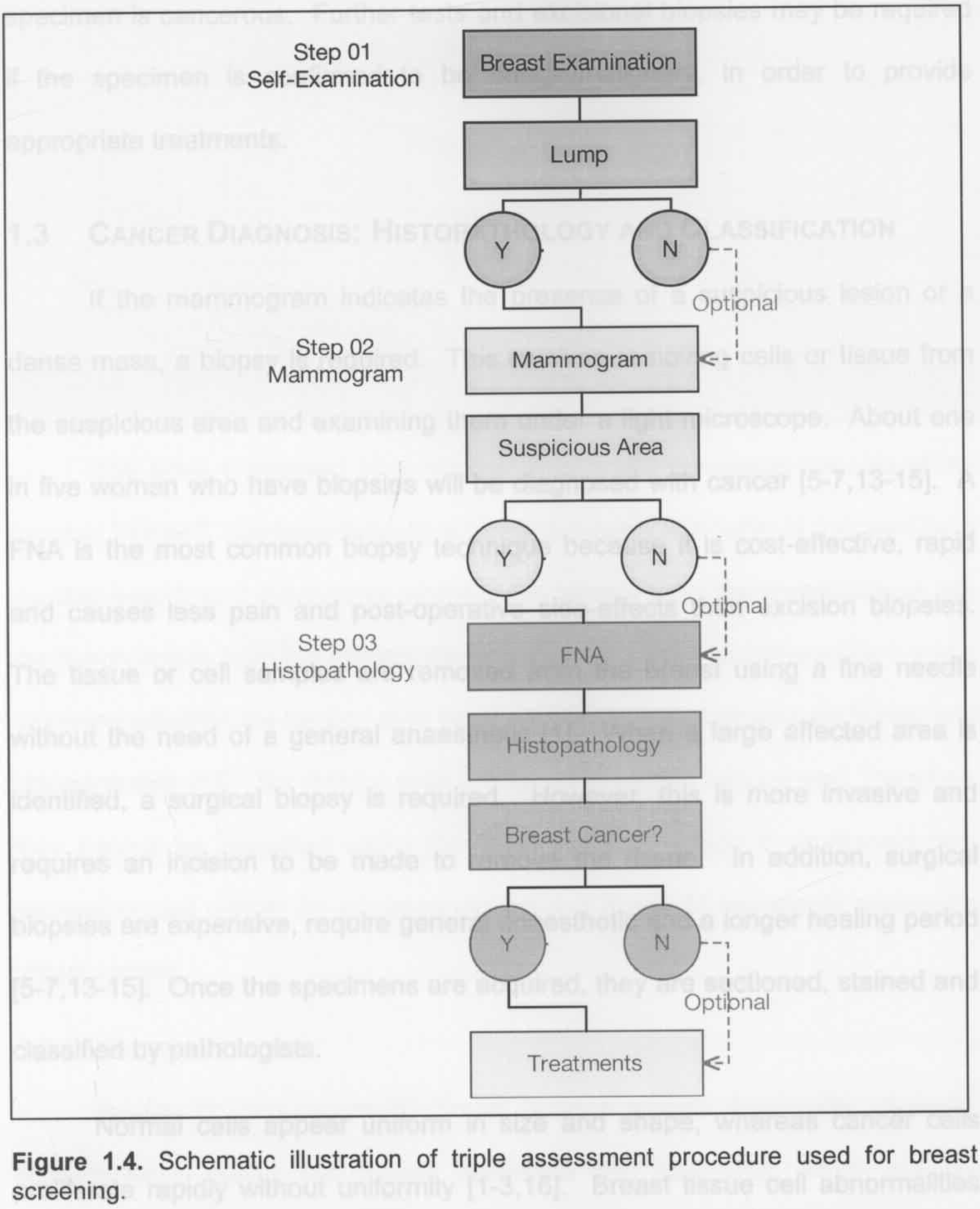


Figure 1.4. Schematic illustration of triple assessment procedure used for breast screening.

The Australian Government provides free annual mammogram screening tests for women between 50 and 69 years of age, the high-risk age group [11]. This program has reduced mortality rates by 25 to 30% in asymptomatic women aged between 50 and 69 [11]. When a suspicious mass is identified, a fine-needle aspiration cytology is performed and the biopsy is examined by a pathologist using a light microscope. The pathology report will indicate if the specimen is cancerous. Further tests and excisional biopsies may be required if the specimen is confirmed to be benign/malignant, in order to provide appropriate treatments.

1.3 CANCER DIAGNOSIS: HISTOPATHOLOGY AND CLASSIFICATION

If the mammogram indicates the presence of a suspicious lesion or a dense mass, a biopsy is required. This involves removing cells or tissue from the suspicious area and examining them under a light microscope. About one in five women who have biopsies will be diagnosed with cancer [5-7,13-15]. A FNA is the most common biopsy technique because it is cost-effective, rapid and causes less pain and post-operative side-effects than excision biopsies. The tissue or cell samples are removed from the breast using a fine needle without the need of a general anaesthetic [1]. When a large affected area is identified, a surgical biopsy is required. However, this is more invasive and requires an incision to be made to remove the tissue. In addition, surgical biopsies are expensive, require general anaesthetic and a longer healing period [5-7,13-15]. Once the specimens are acquired, they are sectioned, stained and classified by pathologists.

Normal cells appear uniform in size and shape, whereas cancer cells proliferate rapidly without uniformity [1-3,16]. Breast tissue cell abnormalities are classified into three main categories: benign; carcinoma *in-situ* and invasive carcinoma (malignant). The tumour size, palpable nodes and metastasis (TNM) stage (Table 1.1) are commonly used to classify breast cancer [1,6].

Table 1.1. TNM staging system for breast cancer [17].

Stage	Tumour (T)	Node (N)	Metastasis (M)
Stage 0	T_{is}	N_0	M_0
Stage I	T_1	N_0	M_0
Stage IIA	T_0	N_1	M_0
	T_1	N_1	M_0
	T_2	N_0	M_0
Stage IIB	T_2	N_1	M_0
	T_3	N_0	M_0
Stage IIIA	T_0	N_2	M_0
	T_1	N_2	M_0
	T_2	N_2	M_0
	T_3	N_1, N_2	M_0
Stage IIIB	T_4	Any N	M_0
Stage IIIC	Any T	N_3	M_0
Stage IV	Any T	Any N	M_1

The letter T followed by a number from 0 to 4 measures the size of the tumour, higher T numbers indicate a larger tumour. The letter N followed by a number from 0 to 3 determines whether the tumour cells are metastasised into the lymph nodes from the primary tumour site. The letter M followed by 0 and 1 assesses whether the tumour cells have spread to distant organs via lymph and blood vessels [17]. No tumour cells are found in benign lesions and, therefore, they cannot be classified using the TNM system. Stage 0 represents pre-invasive breast cancer (*in-situ* carcinomas) and Stage IV represents metastatic breast cancer.

1.3.1 BENIGN LESIONS

Benign or non-neoplastic breast lesions include fibrocystic changes, fibroadenoma, ductal and tubular adenoma and papilloma, which are not cancerous [1,3,16]. If not diagnosed and treated, benign neoplasms may develop into carcinoma *in-situ* and then possibly metastasise to form invasive breast cancer (malignancy) [3].

Fibrocystic changes result from increased ductal epithelia proliferation, which produce a dominant collagenous component in both the epithelial tissue and stroma in the form of ill-defined thickening or lumps [1,3,16]. The formation of a cyst, which can vary in size and density, is the most common fibrocystic change. These changes are related to oestrogen levels and are normally found in women over the age of 20 until menopause [3]. The cytological features are very similar to those of the benign condition fibroadenoma [1,3,16]. Other fibrocystic changes include ductal hyperplasia and adenosis [1].

Fibroadenóma, normally found in younger women between 30 and 40 years of age, is a benign neoplasm that develops in the terminal duct-lobular unit (TDLU) [1,3,16]. It is influenced by the level of oestrogen and is related to the use of oral contraception taken before the age of 20 [1]. The small (< 3 cm^3), firm, round and mobile masses are easily detected by mammography. They are composed of an epithelial component and a substantial stromal neoplastic element both of which display increased levels of collagen deposition [1,3,16].

Intraduct papillomas arise from the duct epithelium and are mostly found in large ducts, such as lacterferious ducts (Figure 1.1) [1,3,16]. The duct contains large complex tumour cells that may be misdiagnosed as malignant. Intraduct papillomas have a higher risk of developing into *in-situ* carcinomas and invasive carcinomas compared with fibrocystic changes and fibroadenoma [1,3,16], therefore, it is important to be able to distinguish amongst the different forms of benign lesions.

1.3.2 *IN-SITU* CARCINOMAS

In-situ breast carcinomas include: ductal carcinoma *in-situ* (DCIS) and lobular carcinoma *in-situ* (LCIS), where the epithelial cells are malignant but have not invaded the basement membrane of the duct or lobule to become invasive carcinoma. DCIS and LCIS are the transitional phases in the evolution of invasive malignancy and, if left untreated, the risk of developing invasive breast cancer increases [1,3,6,16].

Ductal carcinoma *in-situ* is an abnormal proliferation of cells that are confined within the ducts. About 1200 women (10% of all breast cancer types) are diagnosed with DCIS each year in Australia [5,9]. DCIS can be detected as micro-calcifications in a mammogram and classification of the lesion is typically based on the architecture. There are five DCIS subtypes: papillary, micropapillary, cribriform, solid, and comedo. Current diagnosis is still challenging due to the heterogenous nature of tumour growth and the lack of a standardised diagnostic classification system [1,2,6,13,14,16,18-20].

Necrosis and microinvasion of the basement membrane are usually found in the comedo subtype (high-grade DCIS), which has a higher risk of developing invasive breast cancer [18,20]. High-grade DCIS also displays linear branching or coarse granular calcification, formation of abnormal nucleoli and chromatin and clusters of large pleomorphic (vary in size, shape and structure) malignant cells. A low-grade DCIS, such as cribriform DCIS, often shows fine granular calcification, similar to benign lobular calcification, the absence of tubules, stromal infiltration by tumour cells and clusters of mildly pleomorphic malignant cells [1,16].

Lobular carcinoma *in-situ* is confined within the lobules. It is a non-invasive multicentric proliferation of atypical epithelial cells in the lobules and terminal ducts of the breast [1,16,21]. LCIS accounts for about 1% of all breast biopsies and 5% of malignant breast tumours. There are no clinical or mammographic manifestations of LCIS, it carries a higher risk of developing breast carcinoma compared with DCIS, and is often diagnosed together with invasive breast cancer [1,16,21].

1.3.3 INVASIVE BREAST CARCINOMA AND DISTANT METASTASIS

Invasive or malignant breast carcinoma occurs when malignant cells invade beyond the basement membrane of the duct or lobule to the surrounding stromal tissues [1,6,16]. There are two classes of invasive breast carcinoma: invasive ductal carcinoma (IDC) and invasive lobular carcinoma (ILC). The cytological features of IDCs include cell dissociation, nuclear enlargement and irregular nuclear margins, abnormal chromatin patterns, nucleoli, mitoses, and intranuclear inclusions [1,16]. Cytological features of ILC include the presence of signet ring cells, abnormal nuclei and nucleoli, cells that are usually small with eccentric nuclei, intracytoplasmic vacuoles with targetoid mucin, and large cells with abundant cytoplasm in a pleomorphic lobular subtype [1,16]. The tumour cells in both types of invasive carcinomas can develop their own blood vessels to supply oxygen and nutrients [1].

Metastasis is the term used to describe the spread of cancer; the tumour cells can spread through blood or lymphatic vessels to other parts of the body. When the tumour cells are contained within the breast and armpit area, it is termed early breast cancer. Locally advanced breast cancer occurs when the tumour cells spread to places near the breast, such as the chest, but no other parts of the body, whereas secondary cancer is found in other parts of the body and is referred to as metastatic breast cancer [1,16].

1.4 BREAST CANCER TREATMENTS

Treatment of breast cancer varies depending upon the grade and stage of the cancer. Localised carcinoma can be removed completely with a lumpectomy, which contains margins free of tumour. This is termed breast-conserving surgery (BCS) [1,6,7]. At the same time, sampling of axillary lymph nodes is undertaken to determine if lymph node metastases are present. Tamoxifen is used as the primary drug treatment for breast cancer, followed by breast-conserving surgery, radiotherapy, chemotherapy and/or mastectomy [1,6,7]. There is no cure for metastatic breast cancer but treatments are aimed at improving the quality of life by reducing any further spread and growth of cancer [1,6,7].

1.5 ALTERNATIVE SCREENING METHODS

The early and accurate diagnosis of cancer is the single most important factor to improve cancer survival rates and optimise health outcomes for patients. Although the triple assessment is the standard screening protocol for breast cancer, mammography is not highly sensitive for women under 50 due to the constant changes of the breast content during the menstrual cycle. The mammographic results are further complicated by the composition of connective and adipose tissues, as breast tissue is less dense in pre-menopausal women and thus more difficult to view [12]. In addition, histopathology has not changed significantly in the past 100 years. There is a large subjective element in terms of the experience of the pathologist and the sampling procedures used to obtain biopsies, which leads to many false positives and negatives [16,20,22-24]. Further tests may be required to improve the diagnosis and such procedures significantly increase the psychological, physical and financial trauma for patients [7,14].

The recognition of the importance of improved diagnostics in optimising health outcomes has led to the development of a large range of new scanning, imaging and spectroscopic diagnostics. While *in vivo* scanning (computed tomographic scanning, CT) and imaging (magnetic resonance imaging, MRI) techniques have had considerable success in locating tumours, they are also expensive, which has placed limits on their more general use. Therefore, there is an ongoing requirement to develop new diagnostics that are fast, accurate, objective and relatively inexpensive as an alternative to histopathology. Recent research using magnetic resonance spectroscopy (MRS) and vibrational spectroscopy has demonstrated the potential of these techniques to provide alternative and more objective diagnostic tools in cancer research [25-29]. These spectroscopic techniques are used to investigate changes that occur to tissues at a molecular level and subsequently to identify any cellular abnormalities within the breast.

1.5.1 MAGNETIC RESONANCE SPECTROSCOPY

Magnetic resonance spectroscopy, in particular ^{1}H MRS, is used to investigate the biochemical compositions of tissues. This involves measuring the ^{1}H spectrum of the tissue sample, typically obtained from fine-needle aspiration, using water suppression to reduce the very intense signal from the water in the tissues. The resulting spectrum is very complex because of overlapping signals from different biomolecules, however, it is possible to target regions of interest that are diagnostic of disease states [28-32]. As such, MRS is now finding its way into the clinic as an accepted means of diagnosing a variety of diseases including cancerous and precancerous conditions [28].

Human breast tissue is one of the most difficult samples from which to obtain a reliable diagnosis for malignancy using MRS due to signal interference from the adipose tissue and surrounding muscles. Previous research in collaboration with Mountford *et al.* from The Institute for Magnetic Resonance Research (IMRR) has shown that elevated levels of choline (Cho) and its derivatives, phosphocholine and phosphatidylcholine, are associated with malignant breast lesions [28,29,31,33-39]. The choline metabolites result in a significant increase in the relative integrated intensity of the resonance at 3.25 ppm in water-suppressed ^{1}H MR spectra of malignant breast tumours (Figure 1.5) [28,29,31,33-39]. The creatine (Cr), phosphocreatine and lysine resonance at 3.05 ppm is used as an internal standard due to the relatively invariant amounts of these biochemicals in breast tissues. When choline is standardised to creatine, phosphocreatine and lysine (choline-to-creatine ratio), a consistent elevation of the choline level was observed in the malignant tissues (Figure 1.5B) compared with those in normal and benign breast tissues (Figure 1.5A) [28,29,31,33-39]. The ^{1}H MRS results have shown that malignant tissues were distinguished from benign lesions with sensitivity and specificity of 95% and 96%, respectively [28,29,31,33-38].

As previously stated, one of the challenges for MRS is its high sensitivity to the water and lipid resonances. Adipose tissues have strong CH_2 signals from the long-hydrocarbon chain of the lipids, which mask diagnostic information in the ^{1}H MR spectra (Figure 1.6) [29]. Fine needle aspirated biopsies are usually used in MRS studies to reduce the amount of adipose tissue content collected during the biopsy process and to limit sample heterogeneity [29-32,39-41]. However, FNA is usually not as accurate as open-core biopsy, hence, FNA may give rise to false positive and negative results. Furthermore, research has indicated cross-section intervals as small as 100 μm must be examined in order to be accurately correlated with the MRS data [28].

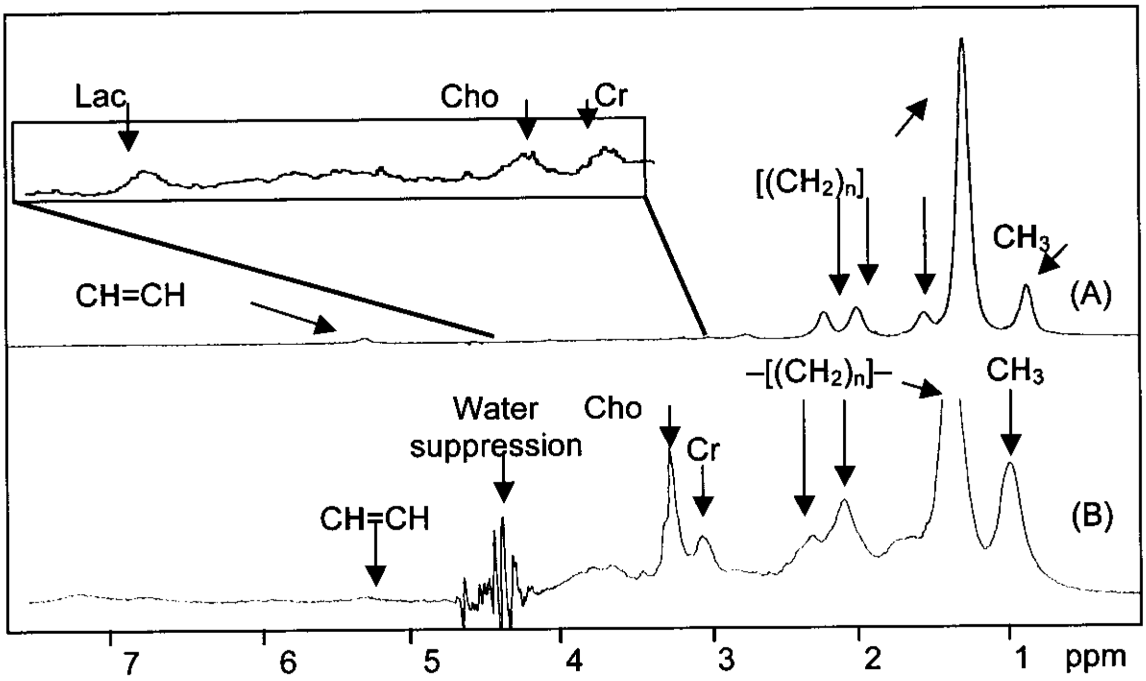


Figure 1.5. *In-vitro* ^{1}H MRS comparison of: (A) a benign breast lesion; and (B) a malignant breast lesion, showing the major differences in choline (Cho) and creatine (Cr) levels [39]. Lac: lactate.

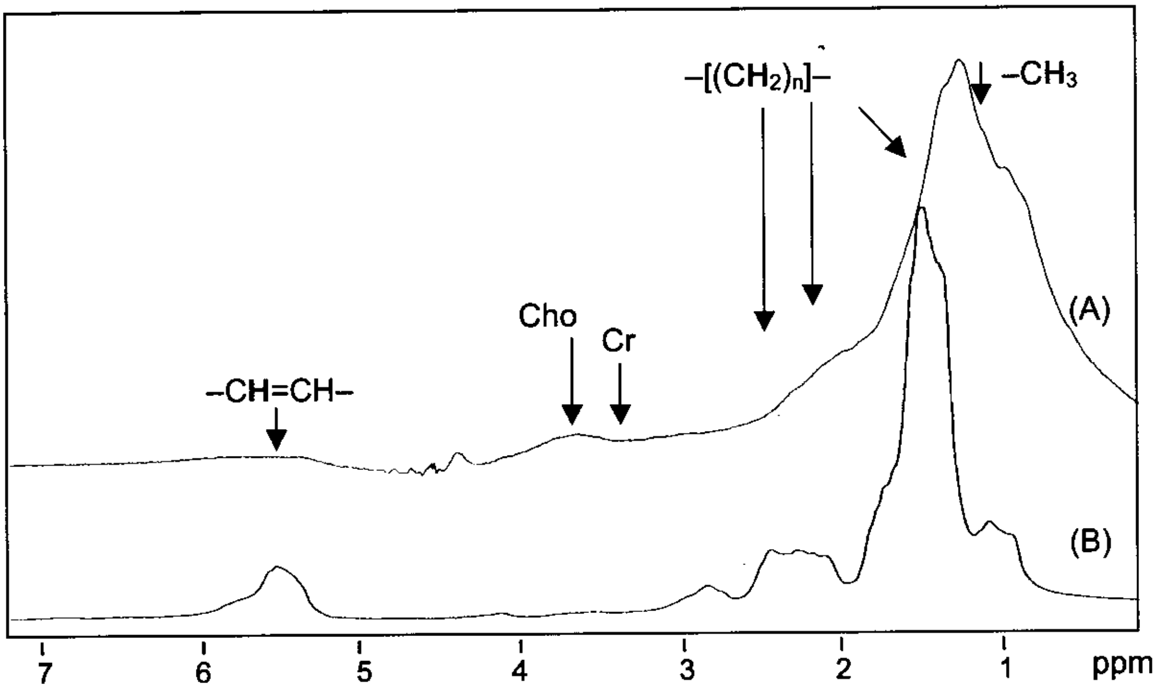


Figure 1.6. ^{1}H MR spectra of lesions with high adipose content masking the diagnostic bands of choline and creatine: (A) benign; and (B) malignant breast lesions [39].

1.6 VIBRATIONAL SPECTROSCOPY IN CLINICAL ANALYSIS

1.6.1 THEORY

Vibrational spectroscopy, infrared (IR) and Raman spectroscopies, has become more routinely used for cancer investigation by various research groups around the world [42-48]. This is ascribed to the development of advanced imaging accessories and relatively low-cost instrumentation compared with the cost of magnetic resonance instruments [49]. Vibrational spectroscopy is a non-destructive technique, allowing a sample to undergo further investigation after vibrational spectral analysis; although this requires judicious choice of laser power by the user for Raman spectroscopy. The mechanism of Raman scattering is different from that of infrared absorption, and Raman and IR spectra provide complementary information [50,51]. A schematic illustration of infrared absorption and the Raman scattering processes is presented in Figure 1.7.

Infrared spectroscopy (Figure 1.7A) is a one-photon effect and the absorption results in a vibrational motion of a molecule [52-54]. Raman spectroscopy is a two-photon scattering process with one photon (from a high energy monochromatic source such as a laser) interacting with a molecule to produce either a virtual state or an electronically excited molecule, which then relaxes back to the ground electronic state [50-55]. The majority of scattered light is unchanged in frequency and the radiation is termed Rayleigh scattering (Figure 1.7B), which is an elastic scattering process [50]. The inelastic process, during which vibrational energy is exchanged, is called Raman scattering [50,51]. The Raman scattered light energy is shifted from that of the incident light by the energies of molecular vibrations.

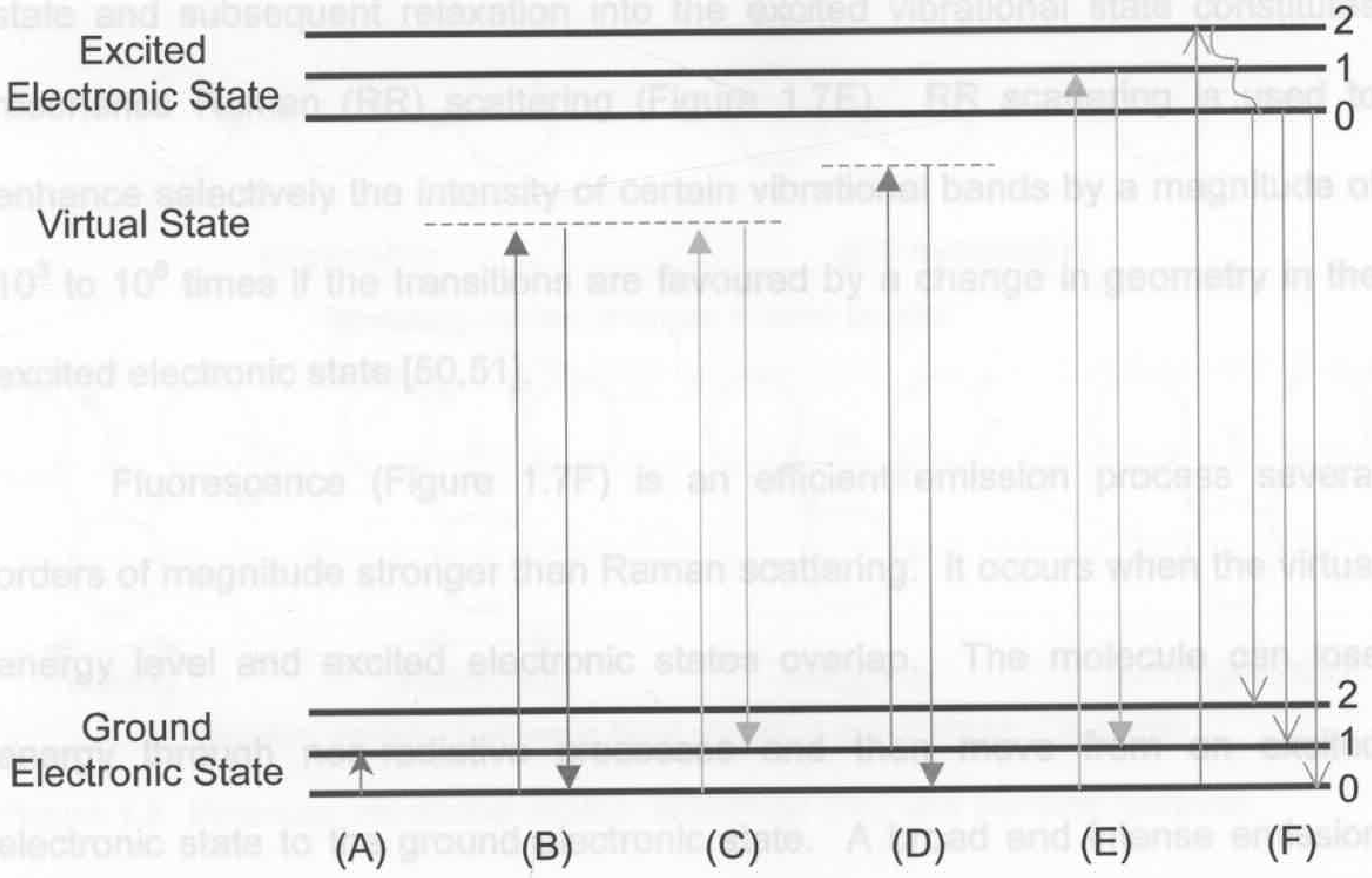


Figure 1.7. Schematic representations of: (A) infrared absorption; (B) Rayleigh scattering; (C) Stokes Raman scattering; (D) anti-Stokes Raman scattering; (E) resonance Raman scattering; and (F) fluorescence. The numbers 0-2 represent different vibrational levels (ν_{vib}) within each electronic state.

Stokes Raman scattering ($\nu_0 - \nu_{vib}$, Figure 1.7C) occurs when the molecule is excited from the ground electronic state to a virtual state and then relaxes into a higher vibrational level of the electronic ground state. Anti-Stokes Raman scattering ($\nu_0 + \nu_{vib}$, Figure 1.7D) occurs when the molecule is initially in the excited vibrational state, and relaxes into the vibrational ground state. These two processes, however, are very weak with 1 in 10^6 photons being Raman scattered as the molecule is promoted to a virtual state. Stokes Raman spectroscopy is more commonly studied than anti-Stokes Raman spectroscopy due to the low population of the excited vibrational states from which the transitions originate [50,51] and, hence, the latter is not used for diagnostics.

When the incident photon has an energy that coincides with the energy of an allowed electronic transition state, the excitation to the excited electronic state and subsequent relaxation into the excited vibrational state constitutes resonance Raman (RR) scattering (Figure 1.7E). RR scattering is used to enhance selectively the intensity of certain vibrational bands by a magnitude of 10^3 to 10^6 times if the transitions are favoured by a change in geometry in the excited electronic state [50,51].

Fluorescence (Figure 1.7F) is an efficient emission process several orders of magnitude stronger than Raman scattering. It occurs when the virtual energy level and excited electronic states overlap. The molecule can lose energy through non-radiative processes and then move from an excited electronic state to the ground electronic state. A broad and intense emission band is produced, which masks the weaker Raman bands. The likelihood of fluorescence occuring increases if the lase excitation used is within the visible region (i.e. towards the UV, higher energy), whereas using lower energy NIR laser excitation will general minimise the fluorescence. [50-55]. The NIR excitation used in Fourier-transform (FT) Raman systems reduces the effect of fluorescence, but also reduces the Raman scattering cross-section (ν^4 law), which decreases with an increases in excitation wavelength [52].

1.6.2 VIBRATIONAL MODES

There are bands due to two common types of vibrational modes in the mid-infrared (MIR) region that are useful for biological studies (4000–650 cm^{-1}): stretching and bending modes, see Figure 1.8. Stretching modes involve changes in bond lengths, either symmetric (ν_s) or anti-symmetric (ν_{as}). Bending, or deformation modes involve changes in bond angles including in-plane rocking (ρ), in-plane scissoring (δ), out-of-plane wagging (ω) and out-of-plane twisting (τ).

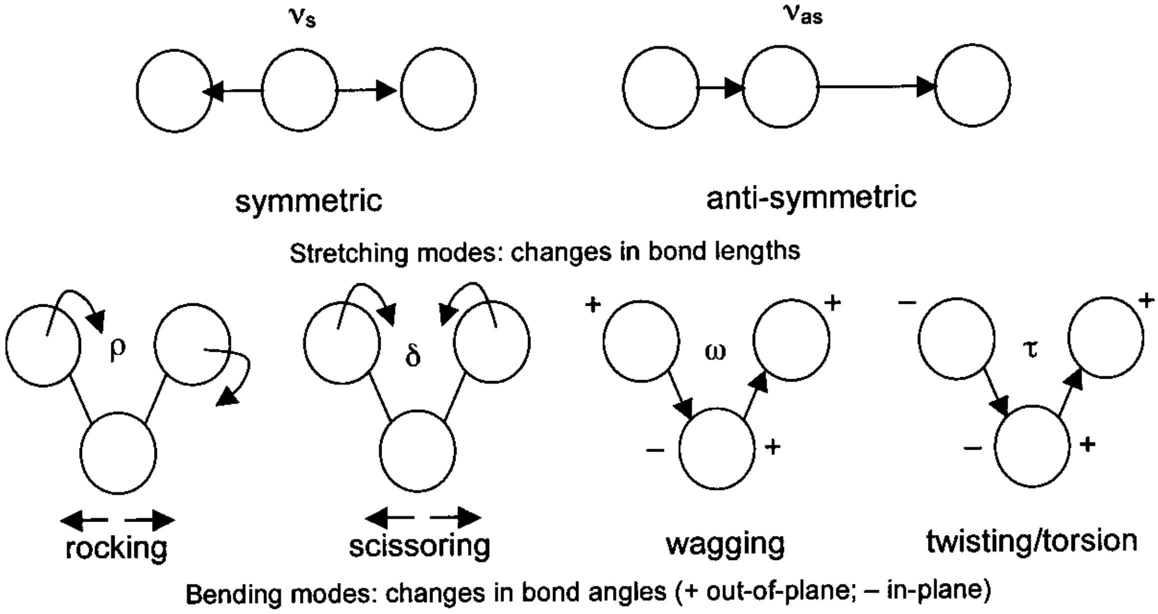


Figure 1.8. Molecular vibrational modes: stretching (top) and bending (bottom).

1.6.3 MEDICAL DIAGNOSTIC APPLICATIONS USING IR SPECTROSCOPY

In the past two decades Fourier transform infrared (FTIR) spectroscopy has shown significant promise for routine medical diagnosis; this is attributable to advances in instrumentation and user-friendly multivariate statistical analysis software. Numerous studies have been undertaken to investigate various human cancers, including breast [33,39,46,47,56-63] , colon [64,65], prostate [66], brain [67-73]; cervix [48,74,75], and other organs [76,77]. The results of these studies consistently demonstrated that FTIR spectroscopy could differentiate malignant lesions from normal and benign tissues.

In some aspects infrared spectroscopy is a simple technique for tissue but there is always the problem of the water to contend with. Tissues can be sectioned and mounted on infrared transparent windows (CaF_2 or BaF_2), or Ag/SnO_2 transflective slides (transmissive to visible light but infrared reflective) without staining prior to analysis [42-47]. Immediately after spectral collection, these sections can be stained and examined, allowing direct correlations with standard histopathological assessments [74,78]. Determination of the disease states are based on changes in the intensities, positions and lineshapes of diagnostic bands (e.g., water, proteins, nucleic acids, lipids and carbohydrates) in the mid-infrared region (4000–650 cm^{-1}) [79]. For example, the amide I band originates mainly from the stretching vibration of the C=O group within the peptide bond, and the intensity, position and lineshape of this band is sensitive to the secondary structure conformation (i.e., α-helix, β-pleated sheets, β-turns and random coil) [54].

Characterisation and differentiation of biological tissues has been investigated using three spectroscopic techniques: single-point spectroscopy, mapping, and imaging, or some combination of these techniques [33,39,46,47,58-62,80,81].

1.6.3.1 SPATIAL RESOLUTION

Spatial resolution is the smallest sample area that IR data could obtain [82]. Both single-point FTIR spectroscopy and synchrotron-radiation-based FTIR microspectroscopy utilise knife-edge apertures for data collection, while FTIR imaging uses a focal-plane array instead of aperture.

The typical aperture for conventional FTIR microspectroscopy is approximately 5 × 5 mm^2 when it is fully opened, which produce the maximum IR illumination on the sample resulting in better IR signal-to-noise ratio (S/N). The dimensions of the breast ducts and lobules ranged from 20 μm to over 100 μm in diameter, however, by reducing the aperture size to the smaller spatial areas (100 × 100 μm^2) yielded poorer S/N [82]. Synchrotron-radiation-based FTIR microspectroscopy (SR-FTIR) offers 100 to 1000 times brighter light source than globar used in conventional FTIR spectrometers, which could resolve the problem of poor S/N in small spatial resolution. The high brightness of the synchtrotron source could investigate small areas with acceptable or better S/N [82-88]. The aperture could be reduced to the spatial resolution of ~ $3 \times 3\ \mu m^2$, which could collect IR information at sub-cellular level, such as cell differentiation [85].

Aperture is not used in the focal-plane array FTIR imaging, instead, multi-channel array detector is used, which comprises individual pixel grid at a special arrangement (e.g., 1 × 16 or 64 × 64) [82]. Each grid has a theoretical spatial resolution of 6.25 μm and the system can collect IR spectra simultaneously in a relatively short time, c.f. IR mapping (both single-point mapping and SR-FTIR mapping).

1.6.3.2 SINGLE-POINT MICROSPECTROSCOPY

Cell proliferation is a characteristic of breast cancer and originates from the lobules or the ducts. This proliferation results in higher nucleic activity than in normal tissues, and the presence of such activity is a proposed method for differentiation of different breast lesions by measuring the absorbances attributed to nucleic acids and proteins. This has been illustrated by the extensive work of Mantsch and co-workers who have successfully distinguished between malignant, normal and benign breast tissues [60,80,89-94]. This earlier research was undertaken with single-point IR microspectroscopy. With this technique, a transmission or a reflection spectrum is collected at a defined area of the tissue section and interpreted by examining spectral differences in diagnostic bands, including those due to proteins (amides), nucleic acids (phosphodiester backbone), lipids ($\delta(CH_2)$) and carbohydrates (glycogen) [79].

However, spectra of biological samples are notoriously complicated due to the heterogenous nature of the sample. As the aperture used in FTIR microspectroscopy is typically large, IR information could include the surrounding tissues, e.g., adipose tissue and cancerous cells, which could affect the IR spectrum and classification of the tissues. In general, there are numerous spectral regions, which contain broad and overlapping bands resulting from the biological components having similar functionality. The spectra of breast tissues are dominated by the epithelial cells of the ducts and lobules, adipose tissues and collagen from the connective tissues [60,80,89-94].

The 3050–2800 cm^{-1} and 1500–900 cm^{-1} spectral regions are particularly informative. The 3600–3100 cm^{-1} spectral region is usually neglected due to the presence of strongly absorbing water bands in the biological samples [80,93]. However, Gao and Ci have also discussed the importance of this region, which contains bands arising from N–H groups in proteins, and $-NH_2$ groups in nucleic acids [59]. The spectra of normal, hyperplastic, benign and malignant breast tissues were found to differ in the relative band intensities of proteins and nucleic acids, together with features assigned to strongly hydrogen-bonded water (SBW) and weakly hydrogen-bonded water (WBW). An increase in the intensity of the SBW band was observed in normal tissues compared with abnormal tissues and a decrease in the intensity of WBW band was observed in benign and malignant tissues. The results, however, could not be used to separate the spectra of hyperplasia, fibroadenoma and carcinoma samples [59].

Ci *et al.* demonstrated that normal and diseased breast tissues had strikingly different spectra [61,62,95]. The band at 970 cm^{-1} from the phosphorylated proteins was observed to become sharper and stronger in the carcinoma IR spectra, which suggests that the concentration of the phosphorylated proteins had increased. The collagen bands at 1338, 1280 and 1204 cm^{-1} were weak and broad in the carcinoma tissue spectra, compared to spectra from benign tissue [61,62,95]. Increased peak intensities for the nucleic acid bands at 1240 and 1080 cm^{-1} were also noted to be indicative of malignancy and result from increased cell proliferation. Various band intensity ratios were used to distinguish malignant from normal tissues [61,62,95].

1.6.3.3 ATTENUATED TOTAL REFLECTION FTIR SPECTROSCOPY

Attenuated total reflection (ATR) or internal reflection spectroscopy (IRS) is a versatile (minimal sample preparation) infrared sampling technique, which utilises the phenomenon of total internal reflection. The basic element of an ATR cell is a high refractive index crystal, such as diamond or zinc selenide (ZnSe) [52,96]. If the angle of incidence (θ_i) is greater or equal to the critical angle (θ_c), then internal reflection will occur [52,96]. Equation 1.1 shows the critical angle relationship between the refractive index (RI) of the sample and the ATR crystal material. The composite diamond/ZnSe internal reflection element (IRE) in this example has a refractive index of 2.403 [52,96].

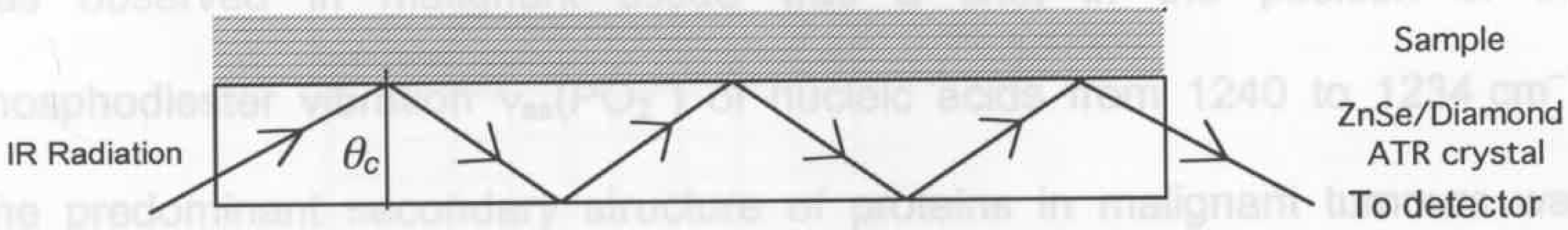


Figure 1.9. Schematic illustration of an attenuated total reflectance cell.

$$\theta_c = \sin^{-1}\frac{\eta_2}{\eta_1} \qquad (1.1)$$

η_2 = refractive index of sample

η_1 = refractive index of the crystal material

The depth of penetration (d_p) is the depth to which the evanescent wave penetrates and, therefore, the depth of sample [52,96]. It is proportional to the wavelength, and is expressed in Equation 1.2:

$$d_p = \frac{\lambda}{2\pi\eta_1\sqrt{\sin^2\theta - \left(\frac{\eta_2}{\eta_1}\right)^2}} \qquad (1.2)$$

d_p = depth of penetration
λ = wavelength (nm)
θ = angle of incidence

According to the literature, the refractive indices η_2 of adipose and glandular breast tissues are approximately 1.455 and 1.4, respectively [97], and if $\theta = 45°$, the RI of ZnSe IRE $\eta_1 = 2.44$, the d_p at 1650 cm^{-1} (amide I) and $\delta(CH_2)$ band of adipose tissue and glandular breast tissues are 1.04 and 0.96 μm for, respectively. The depth of penetration at higher wavenumbers (3000 cm^{-1}, d_p = 0.52 μm for glandular breast tissue) is notably less than at lower wavenumbers (1000 cm^{-1}, d_p = 1.58 μm).

Several research groups have used ATR-FTIR spectroscopy to investigate breast tumours [98-101]. The main characteristic difference that was observed in malignant tissue was a shift in the position of the phosphodiester vibration $\nu_{as}(PO_2^-)$ of nucleic acids from 1240 to 1234 cm^{-1}. The predominant secondary structure of proteins in malignant tumours was reported to be α-helices. The authors also found that the integral area ratios of proteins, lipids and nucleic acids bands, were significantly higher in malignant than benign and normal tissues [98-101]. These findings were indicative of an increasing degree of hydrogen-bonding of oxygen atoms in the nucleic acid backbone in the malignant tissues with decreased concentrations of collagen and glycogen (major components of stroma) [100].

1.6.3.4 FOCAL-PLANE ARRAY FTIR IMAGING

Focal-plane array (FPA) detector used in FTIR imaging collects multiple spectra simultaneously over a relatively large sample area (e.g., image size of 350 × 350 μm^2). The collection time for each image is within a few mintues depending on the spectral resolution compared to point-by-point mapping, where collection time could take tens of hours with a spatial resolution no better than 15 × 15 μm^2 [49]. Each image can be interpreted as an information-rich data cube containing 64 × 64 = 4096 (typical) individual pixel spectra (Figure 1.10).

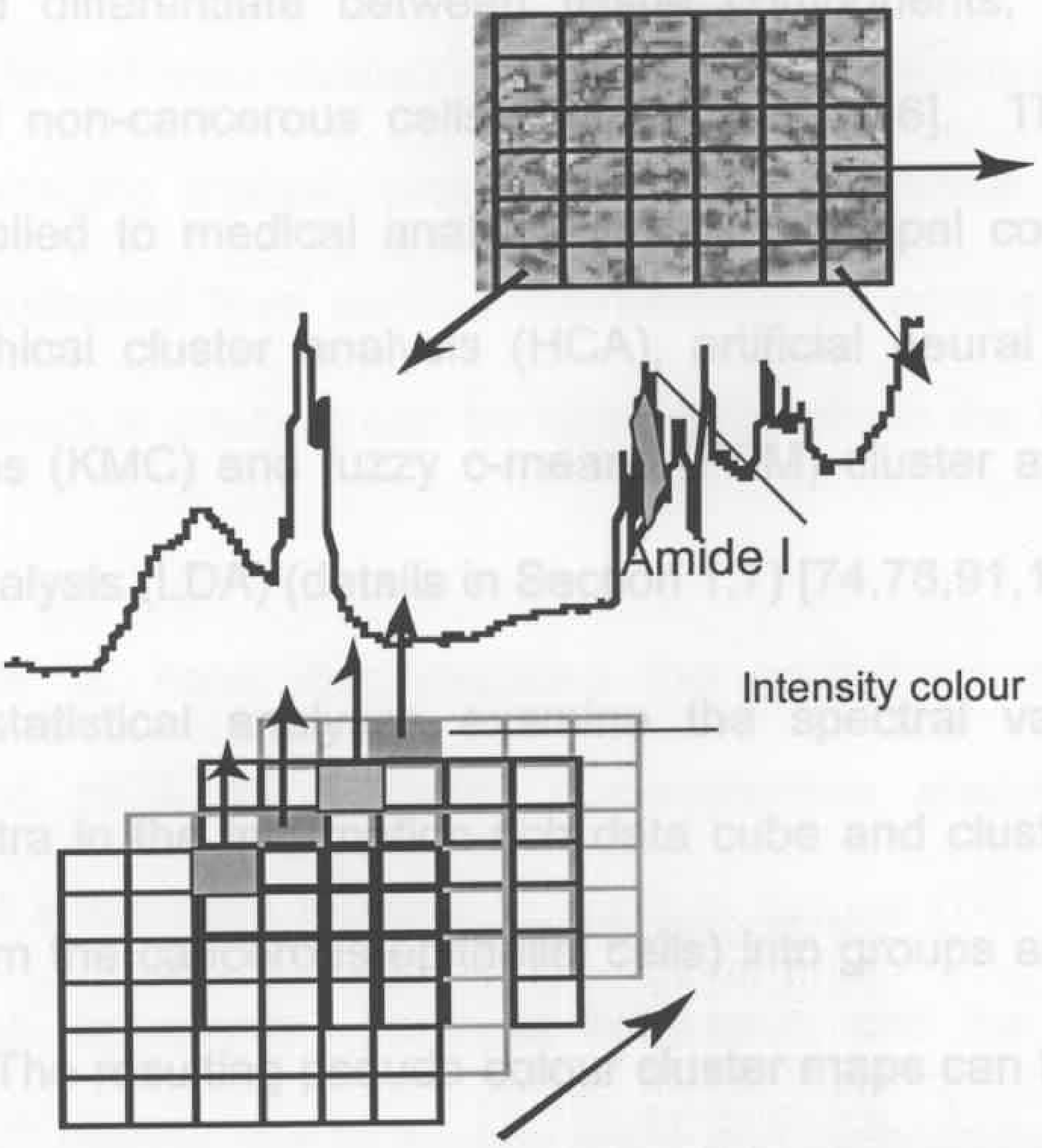


Figure 1.10. Focal-plane array FTIR imaging technique showing multiple grids. Each grid represents an image of the absorbances of the sample at a particular wavenumber. The spectrum of a point on the sample is represented by the line of a particular pixel through all the grids. IR images can be constructed based on the intensity, area, or relative diagnostic band ratios, resulting in a functional group map showing the distribution of specific band(s) of interest on the sample.

Functional group maps, or images, are usually constructed by calculating the area underneath the diagnostic peaks for each spectrum and then plotting this value versus the *x,y* position where the spectrum was collected [49]. These maps or images then illustrate the distribution of each biochemical component (i.e., proteins, nucleic acids and lipids) and can be directly compared with histopathology.

1.6.3.4.1 FPA-IMAGING FOR MEDICAL ANALYSIS

Infrared imaging has become a popular technique for cancer investigation due to its rapid spectral collection and ability to visualise the cellular component distribution in tissue sections. Several authors have employed multivariate statistical analysis (MSA) to complement IR imaging and to identify and differentiate between tissue components, for example, the cancerous and non-cancerous cells [74,78,91,102-106]. The common MSA techniques applied to medical analysis include principal component analysis (PCA), hierarchical cluster analysis (HCA), artificial neural network analysis (ANN), k-means (KMC) and fuzzy c-means (FCM) cluster analysis and linear discriminant analysis (LDA) (details in Section 1.7) [74,78,91,102-106].

These statistical analyses examine the spectral variations amongst individual spectra in the information-rich data cube and cluster similar spectra (e.g., those from the cancerous epithelial cells) into groups and assign specific colour codes. The resulting pseudo-colour cluster maps can then be correlated with the haematoxylin and eosin (H&E) stained sections. As H&E stained sections might not differentiate specific types of cancerous cells, these MSA-based functional group maps could provide vital information in assisting tissue classification and ultimately increase diagnostic accuracy [78].

Lasch *et al.* compared three different clustering algorithms (HCA, FCM and KMC) to assemble IR spectral maps from a colorectal adenocarcinoma sections [78]. The aim of this study was to evaluate which cluster analysis produced the best information for tissue component identification. Different number of clusters (2, 4, 6, 8 and 11) were applied in each algorithm on the carcinoma section and the resulting pseudo-colour maps clearly distinguished the tissue types and could be directly compared with the H&E stained section. It was determined that six clusters were able to classify the major morphological tissue components of the colorectal adenocarcinoma when the cluster maps were compared with the H&E stained section [78]. The results showed that HCA proved to be the best algorithm for tissue structure differentiation, but also the most CPU-intensive image methodology (4.5 hr) compared with the FCM (30 min) and KMC (7 min) clustering algorithms. The authors also developed a dedicated IR imaging analysis program – Cytospec, which is compatible with imaging files collected from various instrumentation companies. The built-in multivariate statistical analysis can be applied directly to the imaging data cube [107].

Wood *et al.* have demonstrated the advantage of combined FPA technology and HCA as a rapid, non-subjective analytical tool for the identification of anatomical features in cervical tissues [74]. Functional group maps based on the amide I band, at 1650 cm^{-1}, and the spectral range of 1300–1200 cm^{-1} (bands due to nucleic acids and carbohydrates) were plotted. The maps clearly illustrated the areas with high concentrations of these biological components; however, the specific tissue types could not be differentiated. The histological features of the high-grade cervical dysplasia were differentiated using HCA in the spectral regions of 1800–950 cm^{-1} and 1700–1570 cm^{-1} [74].

Diem *et al.* examined a large lymph node (~5 mm × 1.5 mm) using IR imaging and HCA. The resulting maps (~100 mm^2) were plotted using the fingerprint region (1800–800 cm^{-1}), which clearly separated follicles, paracortex, medullary cord, capsule and fatty tissues [108]. These imaging results strengthened the FPA-FTIR imaging technique as a potential diagnostic tool. The authors also discussed the advantages and possible disadvantages of infrared imaging (dispersion artefacts) and HCA (long computing time for analysis of large IR image maps).

1.6.3.4.2 IR IMAGING FOR BREAST CANCER DIAGNOSIS

IR imaging studies that have investigated breast cancer have been limited [46,47]. As described in the previous section, DCIS is not classified as a malignant cancer but its presence is believed to be an indicator of a higher risk of developing invasive breast cancer [1]. The uncontrollable cell growth within the DCIS duct results in an increase in the number of nuclei and consequently an increase in the intensity of bands attributed to nucleic acids (phosphate backbone of DNA and RNA) in the IR spectrum [54].

Fabian *et al.* examined a breast tissue section containing a DCIS using the traditional point-by-point mapping technique and imaging using a FPA detector to illustrate the advantages of the imaging technique. Imaging enabled the simultaneous collection of multiple spectra (4096) with a spatial resolution near the diffraction limit. The study also demonstrated the added benefit of detecting micro-areas of high nuclear activity, which were not detected using the traditional point-by-point mapping technique [46,47]. Later research from this group demonstrated that cluster-analysis-based images could classify and correlate with the normal, benign, DCIS and malignant lesions [109]. The use of cluster analysis enabled the separation of individual breast tissue components but fundamental changes in the shapes and sizes of nuclei within the breast lesions were not well differentiated [46,47].

K-means cluster (KMC) analysis obtains a hard class membership for each spectrum, and each class membership is represented by 0 or 1 (detailed in Section 1.7.2) [78]. The analysis had been used to differentiate cells from the medium and non-cellular components in cultures of Hs578B human breast normal and cancerous cell lines [110]. Two clusters were found to separate the spectral data into cell and non-cell clusters. The discrimination of normal and carcinoma cells were then classified by more advanced artificial neural network analysis [110].

1.6.3.5 SYNCHROTRON-RADIATION-BASED FTIR MAPPING

FTIR microspectroscopy, using a globar source, is limited by the low intrinsic brightness of the source and diffraction effects that occur when the aperture size is reduced below 35 to 50 µm – depending on the type of infrared microscope [88]. As a consequence, less light enters the detector and there is a decrease in the signal-to-noise (S/N) ratio of the spectrum [88]. Synchrotron-radiation-based FTIR (SR-FTIR) microspectroscopy takes advantage of the extremely bright (1000 times more intense than a conventional globar IR source), and non-divergent characteristics (highly collimated like a laser) of the synchrotron source in which a similar intensity beam of light is focused into a very small area [88]. These characteristics allow the use of significantly smaller aperture sizes (3–10 µm), faster data collection, good spectral S/N, and spatial resolution limited only by diffraction [111]. As such, SR-FTIR spectroscopy is a new approach for the study of biological samples such as applications in cancer diagnosis including those of the prostate, cervix and breast, as well as micro-organism identification [86,88,112-115].

1.6.3.5.1 SR-FTIR SPECTROSCOPY FOR MEDICAL DIAGNOSIS

Wetzel *et al.* investigated collagens and elastin in twenty human coronary tissue sections to show the potential of using SR-FTIR mapping for investigating vascular changes and protein composition associated with cardiovascular disease [116]. The preliminary results showed that different collagen and elastin distributions were located in the fibroatheromas and could be used to identify the potential for cardiac attack [116].

A study of cytokinetic (cell division) and motile (locomotion) prostate cancer cells using SR-FTIR spectroscopy was investigated by Gazi [117]. The biospectroscopic functional group maps of the prostate cancer cells were plotted to illustrate the process of cytokinesis and locomotion. These maps were used to gain an understanding of the fundamental biochemical processes that occur at the mid-body of the cytokinetic cells and at the lamellipodium[1] of motile cells. A reduction of the intensity distribution of the amide I and II bands was observed at the mid-body in the IR functional group maps and was proposed to arise as a result of the actin-myosin ring contractions. The motile PC3 prostate cancer cells also revealed a heterogenous distribution of lipid intensity at the lamellipodium proposed to be due to the process of disarrangement and lateral membrane flow, which are consequences of cell

---

[1] Lamellipodium - a flattened extension of a cell, by which it moves over or adheres to a surface.

migration [117].

Tobin *et al.* have used both SR-FTIR and conventional globar FTIR spectroscopy to examine oral epithelial tumour tissues and cultured cervical epithelial carcinoma cells [112]. Initial comparisons revealed a significant spectral improvement when using SR-FTIR microspectroscopy under identical collection conditions. The tumour cells were distinguished from stroma in the oral tissues using hierarchical cluster anlaysis. Studies on epithelial cells using a growth-stimulating hormone enabled the detection of key changes in the infrared spectra that were related to the activation of the growth factor signalling mechanism [112].

1.6.3.5.2 SR-FTIR SPECTROSCOPY FOR BREAST CANCER ANALYSIS

There are two reports in the literature of the use of SR-FTIR spectroscopy for breast cancer analysis [118,119]. In the first report, Liu and co-workers used single-point SR-FTIR microscopy for an *in-vitro* study of breast tissues. The position of the bands attributed to nucleic acids, at 1082 cm^{-1}, and DNA, at 968 cm^{-1}, were reported to shift in the spectra of benign and malignant tisuse compared to normal tissue. This shift was evident in the spectra that were presented in the paper, however, the authors did not clearly state the original positions of these bands or the magnitude of the shift [118].

The authors also noted the presence of two peaks at 1464 and 1474 cm^{-1} that were observed only in the SR-FTIR spectra of these samples [118]. The intensities of these two bands were relatively higher in malignant lesions than in normal and benign tissues [118]. In addition, a peak at 1500 cm^{-1} assigned to the $\nu(C=C)$ vibration of pyrrole in red blood cells was observed in the malignant tissue. It was hypothesised that this may be due to an increase in red blood cells in the process of vasculogenesis in the breast tissues during their proliferation [118]. However, the authors did not use a mapping technique to compare the different areas of the tissues.

Breast calcifications were investigated using SR-FTIR microscopy by Baker *et al.* [119]. There are two types of micro-calcifications, calcium oxalate dihydrate and calcium hydroxyapatite [1,16]. The former is usually found in the benign lesions whereas the latter one is associated with malignancy, however, both micro-calcifications are normally difficult to identify morphologically. By utilising SR-FTIR spectroscopy and principal component analysis, major differences in the IR area maps could be observed [119].

1.6.4 MEDICAL DIAGNOSTIC APPLICATIONS USING RAMAN SPECTROSCOPY

IR spectroscopy provides fast and reliable disease diagnosis, however, the strong and broad water bands in the 3600–3100 cm^{-1} spectral region mask other bands and subsequently information about amino acids, lipids, nucleic acids and other biological components [58]. Raman spectroscopy, in comparison, is more suitable for the investigation of biological samples, since water is a poor Raman scatterer [52]. Raman analytical techniques for investigation of biological samples were well-established with the introduction of near-infrared Fourier transform (NIR-FT) Raman spectroscopy in the mid-to-late 1980's [120,121]. The Raman effect has a very small quantum yield and when visible excitation is used, the weak Raman bands are usually overlaid by broad fluorescence bands from the cellular constituents [120]. With the choice of laser excitation from the ultraviolet to the near-infrared, Raman spectroscopic applications to breast cancer research have been growing using both macro- and micro-spectroscopic approaches [33,119,122-140].

1.6.4.1 MACROSPECTROSCOPIC APPROACHES

Alfano *et al.* first recorded Raman spectra of human breast tissues using FT-Raman spectroscopy with a laser excitation of 1064 nm [141]. Differences in relative intensities of characteristic bands were observed between normal and malignant lesions. They found that the relative intensities of the bands at 1445 cm^{-1}, due to the $\delta(CH_2/CH_3)$, and 1650 cm^{-1}, assigned to the amide I band of proteins, could be correlated with the disease classification. Changes in ratios of these two bands suggested an increase in protein content as a result of rapid cell growth [136,142,143]. Raman spectroscopic studies of biomaterials in the regions of 3500–2700 cm^{-1} and 1800–400 cm^{-1} contain bands that are due to tissue components including: carbohydrates, nucleic acids, proteins, and lipids [144]. The lineshape and the intensities of the amide I band and the ring breathing mode of phenylalanine, at 1003 cm^{-1}, were relatively stronger in diseased tissue compared to normal tissue [128].

1.6.4.2 MICROSPECTROSCOPIC APPROACHES

The high laser powers and long exposure times used in FT-Raman spectroscopy to acquire high quality spectra are not ideal. Silicon-based charge-coupled device (CCD) detectors have dramatically improved the instrumentation used to measure Raman spectra. Recently, investigators have used ultraviolet (< 300 nm) and near-infrared excitation wavelengths (700 to 850 nm), where fluorescence is significantly reduced and silicon CCDs' are still sensitive to Raman-scattered light [145]. Although background signals are larger at these wavelengths, this background is broad and can be unambiguously removed from the sharp Raman bands to yield useful spectra [130,136-138].

Following the earlier characterisation of human breast tissue using Raman spectroscopy by Alfano's preliinary work, Frank and Redd *et al.* investigated normal, abnormal, benign and malignant tissues using laser wavelengths ranging from 406 to 830 nm [130,136-138]. They reported that the band present at 1439 cm^{-1} in the normal tissue shifted to 1450 cm^{-1} in the infiltrating ductal carcinoma and they attributed this change to increased protein concentrations in malignant samples. The 1654:1439 cm^{-1} band area ratio was found to differentiate IDC from normal tissues, however, fibroadenoma and IDC were not differentiated [130,136-138]. Spectra of formalin-fixed human tissue revealed Raman features assigned to lipids and carotenoids in addition to the typical biological peaks. The best defined lipid features were observed using 782- and 830-nm excitations, while carotenoid features were strongest in the 488–515 nm range due to their resonance enhancement [130,136-138].

Yazdi *et al.* studied normal and malignant cultured breast and cervical cells using resonance Raman spectroscopy with 257-nm excitation [124]. The sensitivity and selectivity of UV-excited resonance Raman spectroscopy illustrated the potential of this technique for studying certain cellular biochemical changes, especially changes in DNA that occur during cancer development. Spectra of cells were found to closely resemble that of DNA, with peaks around 1580, 1480 and 1330 cm^{-1} assigned to the nucleotide bases. In addition to these bands, the uracil base in RNA results in a peak at 1230 cm^{-1}. Strong tryptophan and tyrosine contributions appear in the 1670–1520 cm^{-1} range. The Raman band ratios, 1480:1614 cm^{-1} and 1480:1540 cm^{-1}, are sensitive to the concentration of nucleic acids relative to cellular proteins, since nucleic acid concentrations are higher in malignant than in normal cells. Normal and malignant cells could also be differentiated by using the 1330:1480 cm^{-1} band intensity ratio. The results of this study indicated that there were significant differences in the UV resonance Raman spectra of normal and cancerous cells. These differences were related to changes in nucleotide/protein concentrations in the cell, as well as conformational changes in the nucleic acids associated with the malignant cell phenotype.

Spectral classification of human breast tissues (normal, fibroadenoma and infiltrating ductal carcinoma) by Raman spectroscopy was investigated for comparison with their histological diagnoses [133,142]. Using principal component analysis, Manoharan *et al.* were able to classify 14 of 15 normal, 13 of 15 benign and 31 of 31 malignant breast tissues, while spectra from benign and malignant tissues were not able to be classified when the peak ratio of 1650:1445 cm^{-1} against the Raman shift of $\delta(CH_2)$ were compared [133,134].

Chowdary and co-workers showed that significant spectral diffference were observed in normal, benign and malignant breast tissues using a 785-nm excitation [146]. These differences included the presence of lipid bands in normal tissues and stronger Raman bands attributable to proteins were found in benign and malignant tissues. Principal component analysis successfully distinguished different conidtions in the spectral region of 1750–1400 cm^{-1} [146].

Most of the previous work has focussed on differentiating malignant breast lesions from normal and benign tissues. In order to understand the relationship between morphology and Raman spectral information, Haka *et al.* studied the chemical composition of breast duct micro-calcifications [122]. The results showed that there are two types of calcifications: type I, calcium oxalate dihydrate, and type II, calcium hydroxyapatite. Type I calcification was mainly found in benign lesions whereas calcium hydroxyapatite was found in malignant lesions. Each calcification type has a unique Raman band in the region of 1800–500 cm^{-1}. Using PCA, the authors were able to distinguish micro-calcifications in benign and malignant ducts with a sensitivity of 88% and a specificity of 93% and the results showed that Raman spectroscopy has a higher accuracy for locating micro-calcifications than X-ray mammography [122].

1.6.4.3 RAMAN MAPPING AND IMAGING

Similar to FPA-FTIR imaging, Raman imaging is also gaining popularity due to the development of a new generation of highly sensitive detectors and imaging accessories. There are two methods possible to construct functional group images: (1) single-point mapping, and (2) laser-defocussed filter imaging. The single-point mapping utilises a standard grating to collect Raman spectra (step or raster scan) within the specified area, then functional group maps are constructed based on a particular wavenumber or region. In contrast, laser-defocus imaging employs filters to collect images at a defined Raman band by defocussing (expanding) the laser beam. The collection time of the single-point mapping is usually longer than laser-defocussed filter imaging.

Shafer-Peltier and co-workers combined Raman mapping and multivariate statistical analysis to investigate the subcellular components in various breast diseases [122,126,134,140,147]. A mathematical model was designed based on the Raman spectra of individual tissue components, which included cell cytoplasm, cell nucleus, adipose tissue, β-carotene, collagen, calcium hydroxyapatite, calcium oxalate dihydrate, cholesterol-like lipid deposits, and water. The fitted models demonstrated that in an infiltrating ductal carcinoma, there was a decrease in the adipose content compared to normal and benign lesions. However, the models could not differentiate between the fibrocystic changes and IDC lesions [140]. Raman images were plotted based on the specific morphological elements including collagen, cell cytoplasm and cell nucleus and such images showed good spectral correlations with the adjacent stained sections [140].

Tong *et al.* investigated the characterisation of normal breast duct epithelia using mapping at an excitation wavelength of 840 nm. Different cellular compounds within the epithelial layer, such as nuclei and the cytoplasm, were identified by principal component analysis and k-means cluster analysis of the Raman spectra [125].

1.7 MULTIVARIATE STATISTICAL ANALYSIS - THEORY

As described in Section 1.6.1.3, combining spectroscopic imaging techniques with multivariate statistical analysis, including: principal component analysis; non-hierarchical cluster analysis (k-means cluster and fuzzy c-means cluster analysis); and hierarchical cluster analysis have the potential to become powerful diagnostic techniques. Tissue types and architectures can be easily visualised and interpreted to produce maps based on pre-defined diagnostic regions.

The multivariate statistical analyses, in general, examine individual spectra (including those that are difficult to differentiate by visual examination of the spectra), cluster similar spectra into specific groups (e.g., those from cancerous cells) and assign each group a unique colour code or location [148-152].

Tissue sections examined using the FPA-FTIR imaging and SR-FTIR mapping produce a large spectral data cube that have various spectral background, noise and artefacts (i.e., anomalous dispersion). These could give rise to false results and the data need to be pre-processed, e.g., baseline correction, smoothing, or taking derivatives, prior to applying multivariate statistical analyses.

1.7.1 PRINCIPAL COMPONENT ANALYSIS

Principal component analysis is widely used in spectroscopy to discriminate between different types of human cancers. It aims to simplify a multi-dimensional data set by reducing it to lower dimensions for analysis [150-152]. It transforms the data to a new co-ordinate system such that the greatest variance by any projection of the data comes to lie on the first principal component (PC1) axis, the second greatest variance on the second co-ordinate (PC2), and so on. In most cases, PC1 and PC2 co-ordinates contain the most important aspects of the data derived from the spectral dataset, see Figure 1.11. For spectral interpretation, the PC scores are usually used to separate the major spectral differences among normal, benign and malignant breast lesions. PCA can be applied to both single-point spectroscopy and imaging.

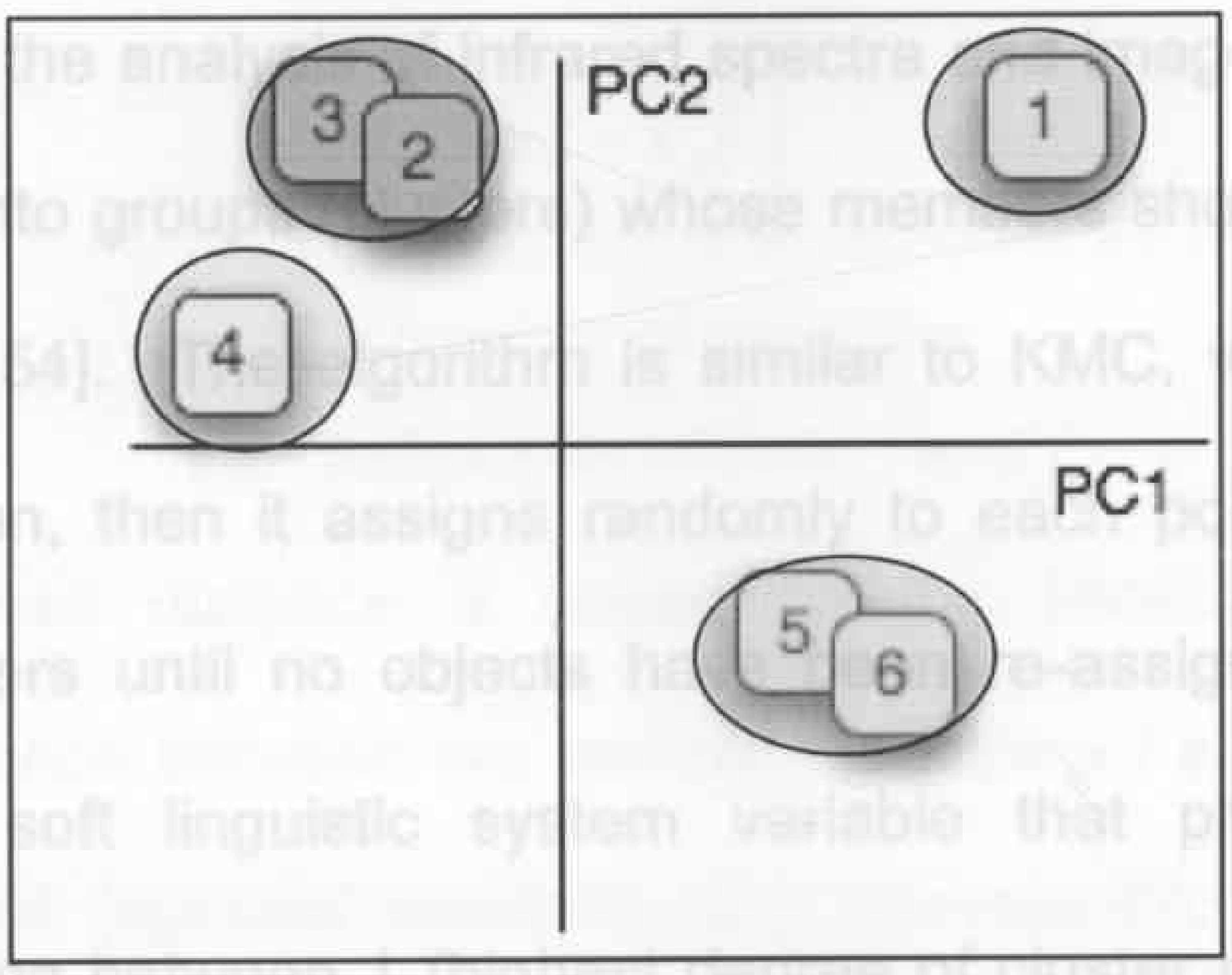


Figure 1.11. Plot of scores of PC2 vs. PC1 showing different groups in specific colours.

1.7.2 K-MEANS CLUSTER ANALYSIS

The k-means algorithm is a non-hierarchical cluster analysis technique commonly used with spectroscopic images. The algorithm clusters objects based on their attributes into k partitions. It is a variant of the expectation-maximisation algorithm in which the goal is to determine the k-means of data generated from gaussian distributions. The objective is to minimise total intra-cluster variance [150,153].

The algorithm starts by partitioning the input points into initial k sets; it then calculates the mean point, or centroid, of each set. It constructs a new partition by associating each point with the closest centroid. Then the centroids are recalculated for the new clusters, and the algorithm is repeated by alternate application of these two steps until convergence, which is obtained when the points no longer switch clusters. The result of KM clustering assigns to each spectrum a hard class membership (either 0 or 1) for each cluster of spectra [153].

1.7.3 FUZZY C-MEANS CLUSTER ANALYSIS

Fuzzy c-means clustering is a non-hierarchical clustering method commonly used in the analysis of infrared spectra and images. This technique partitions objects into groups (clusters) whose members show a certain degree of similarity [107,154]. The algorithm is similar to KMC, where a number of clusters are chosen, then it assigns randomly to each point co-efficients for being in the clusters until no objects have been re-assigned. FCM cluster analysis uses a soft linguistic system variable that produces values of membership function between 1 (highest degree of cluster membership) and 0 (unrelated membership) as opposed to either 0 or 1 in the KMC algorithm. The sum of the values of the C cluster membership for one object equals 1 [107,154]

The advantage of the fuzzy clustering over KM clustering is that both outliers and data that display properties of more than one class can be characterised by assigning non-zero class membership values to several clusters. The membership values of FCM clustering functional group maps can be encoded by color intensities, with a high colour intensity representing higher class membership [107,154].

1.7.4 HIERARCHICAL CLUSTER ANALYSIS

Hierarchical cluster analysis tests the spectra for their similarity to each other. The analysis divides similar spectra into clusters and displays clusters based on the heterogeneity on a dendrogram [148-151].

A spectral region from the data (original or pre-processed) is selected, in order to determine the similarity between objects, and the distance matrix is calculated using Minkowski distance (Equation 1.3):

$$d_{ij} = \left[\sum_{k=1}^{K} \left| x_{ik} - x_{jk} \right|^{p} \right]^{1/p} \qquad (1.3)$$

K = number of variables

i, j – indices for objects *i and j*

The Euclidean distance is based on the Minkowski distance and measures the distance between two samples (samples *i and j*) where *k* is the number of variables (spectral elements, e.g., wavelenth). For example, the Euclidean distance between two objects 1 and 2 is:

$$d_{12} = \sqrt{\left[(x_{11} - x_{21})^2 + (x_{12} - x_{22})^2 \right]} \qquad (1.4)$$

This distance can be extended into mutli-dimensioned space, e.g., a spectrum is a point in a space which has as many dimenisons as the spectra have frequencies. The spectral distance indicates the degree of spectral similarity, i.e., spectra with a spectral distance or heterogeneity of 0 are entirely identical (within the frequency ranges tested) [148-151]. The higher the distance between spectra, the more substantial the spectral difference [148-151].

The calculated distance is then used to link the objects by applying hierarchical (agglomerative) clustering such as Ward's algorithm (Equation 1.5). The algorithm determines the growth of the heterogeneity factor, H, which tries to find the most homogeneous groups possible [151].

$$H(r,i) = \frac{[n(p) + n(i)] \cdot D(p,i) + [n(i) + n(q)] \cdot D(q,i) - n(i) \cdot D(q,i)}{n + n(i)} \qquad (1.5)$$

p and q - individual clusters; r - merged p and q clusters;

D - spectral distance; n(i) is the number of spectra merged in the I cluster

The clustering process repeats itself until all objects merge into larger object groups and the resulting clusters are depicted as a dendrogram. The dendrogram displays the distance of heterogeneity among groups (Figure 1.12); the less heterogeneity in a group, the more similar is the spectral information or group [148-151].

An example containing six cities is compared (Figure 1.12) to determine the distances among them: Hong Kong (1), Sydney (2), Melbourne (3), Brisbane (4), New York (5), and Paris (6). By applying HCA, Sydney (2), Melbourne (3), and Brisbane (4) are clustered in one group as these cities are located in Australia; Hong Kong (1) is closer to Australia and is clustered with the Australian cities, both New York and Paris are further away from Australia but they are closer to each other and therefore, they cluster together.

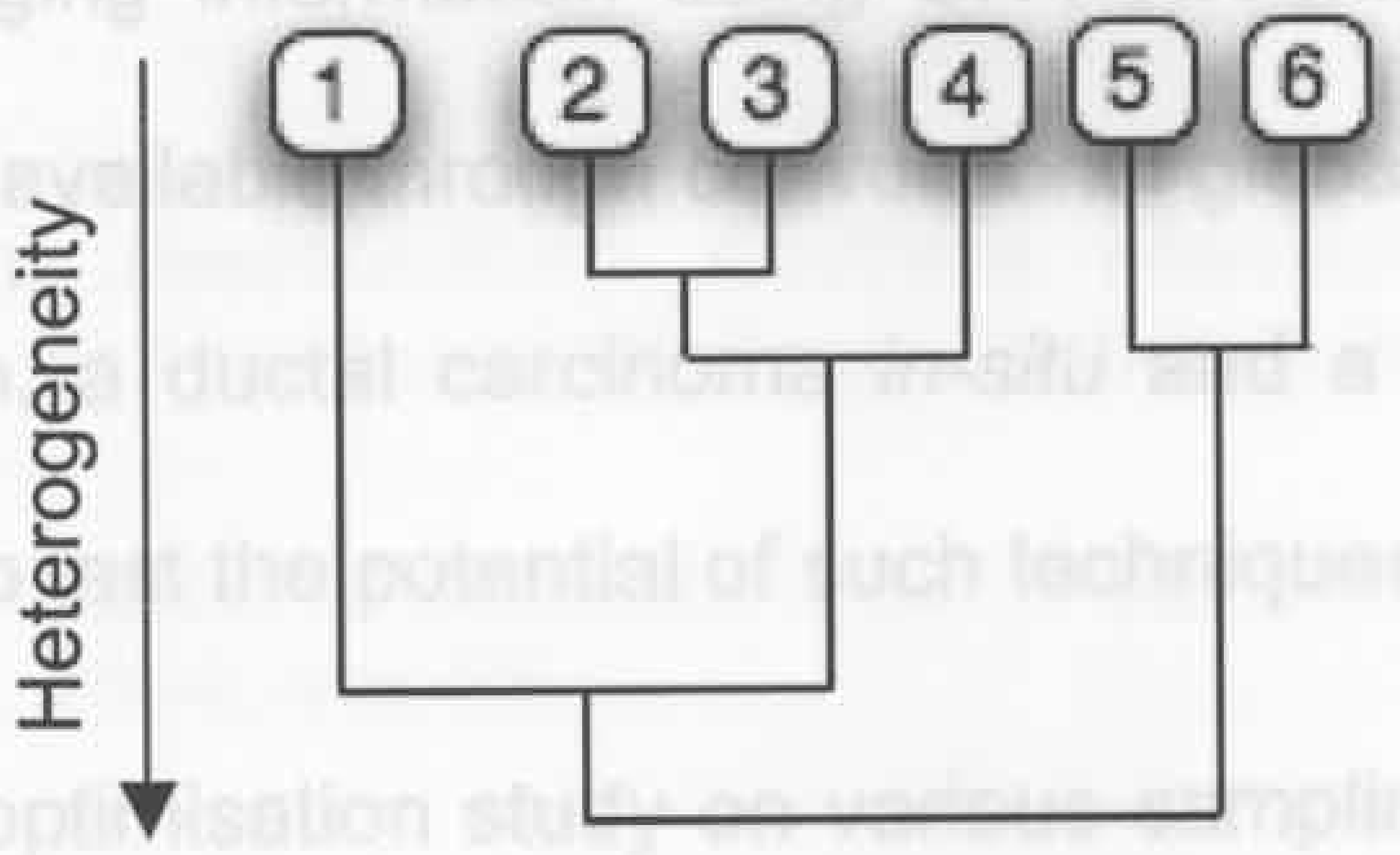


Figure 1.12. A dendrogram for six cities located in the world based on the distance difference: (1) Hong Kong; (2) Sydney; (3) Melbourne; (4) Brisbane; (5) New York; and (6) Paris..

1.8 RESEARCH AIMS

This project primarily aims to develop robust methodologies for rapid and accurate diagnoses of breast cancer based on the abilities of IR and Raman spectroscopic techniques to distinguish biochemical changes that are diagnostic of cancer and pre-malignant conditions. It is envisaged that these could be used to complement current pathological testings using single-point spectroscopy.

The research described implements both IR and Raman imaging techniques to resolve the challenges associated with the heterogenous nature of tumour growth and the lack of a standardised diagnostic classification system of transitional cancers, including DCIS and LCIS. FPA-FTIR imaging combined with multivariate statistical analysis will be explored to differentiate the various cancer stages that can originate from breast ducts. Raman imaging will also be explored as a method to characterise a LCIS breast lesion and to locate any micro-calcifications within the breast ducts.

To date, the application of SR-FTIR spectroscopy for medical analysis has been limited. This research aims to examine what further molecular spectroscopic imaging information using SR-FTIR microspectroscopy can be obtained over that available through conventional globar sources. A preliminary study on a benign, a ductal carcinoma *in-situ* and a malignant breast lesion were undertaken to test the potential of such techniques.

Finally, an optimisation study on various sampling techniques, including: the assessment of contamination during tissue sectioning; and spectroscopic analysis using infrared and Raman spectroscopy, will be discussed. The aim is to develop a systematic workflow that can integrate vibrational spectroscopy seamlessly into current pathological procedures. Ultimately, the research is directed towards the design and development of early and reliable detection of cancerous and pre-cancerous conditions in order to optimise the treatment and reduce the mortality rate resulting from malignant tumours.

1.9 REFERENCES

[1] G. McKee, *Cytopathology of the Breast.* 1st ed., Oxford University Press, Boston, 2002.

[2] R. Blamey, A. Evans, I. Ellis, R. Wilson, *Atlas of Breast Cancer.* 1st ed., Merit Publishing International, Basingstoke, 1994.

[3] R. Fechner, S. Mills, *Breast Pathology: Benign Proliferations, Atypias and in-Situ Carcinoma.* 1st ed., American Society of Clinical Pathologists Press, Chicago, 1990.

[4] National Cancer Institute, Genetics, Causes, Risk Factors, Prevention of Breast Cancer. http://www.cancernet.com 2000.

[5] National Breast Cancer Centre, http://www.breasthealth.com.au 2004.

[6] The Susan G. Komen Breast Cancer Foundation (USA), http://www.komen.org 2005.

[7] Imaginis, http://www.imaginis.com 2005.

[8] International Agency for Research on Cancer, *Software: Globocan 2002,,* 2005.

[9] C. Stevenson, *Breastscreen Australia Monitoring Report 2001-2002.* Australian Institute for Health and Welfare, Canberra, 2005.

[10] I. McDermid, *Cancer in Australia 2001.* Australian Institute of Health and Welfare, Canberra, 2004, Vol. 28.

[11] Australian Institute of Health and Welfare, Cancer Database 2001. http://www.aihw.gov.au/cognos/cgi-bin/ppdscgi.exe?DC=Q&E=/Cancer/canceragerratesv7 2005.

[12] D. Spicer, M. Pike, *Risk Factors in Breast Cancer.* Churchill Livingstone, Philadelphia, 1999, 47-54.

[13] BreastScreen Australia, http://www.breastscreen.org.au 2005.

[14] CancerBACUP, http://www.cancerbacup.org.uk 2005.

[15] Breastdoctor, http://www.breastdoctor.com 2004.

[16] J. Harris, *Diseases of the Breast.* 3rd ed., Lippincott-Raven, Philadelphia,

2004.

[17] F. Greene, D. Page, I. Fleming, A. Fritz, C. Balch, D. Haller, M. Morrow, *Breast in American Joint Committee on Cancer: Ajcc Cancer Staging Manual.* 6th ed., Springer, New York, 2002, 223-240.

[18] H. Burstein, K. Polyak, J. Wong, S. Lester, C. Kaelin, *N. Engl. J. Med.* 350 (2004) 1430-1441.

[19] Imaginis, Ductal Carcinoma *in-Situ.* http://www.imaginis.com/breasthealth/dcis.asp 2005.

[20] K. A. Skinner, M. J. Silverstein, *Endocr. Rev.* 8 (2001) 33-45.

[21] E. Frykberg, *Breast J.* 5 (1999) 287-350.

[22] K. A. Skinner, *The Clinical Management of Ductal Carcinoma in-Situ, Lobular Carcinoma in-Situ and Atypical Hyperplasia of the Breast.* 1st ed., National Breast Cancer Centre, Sydney, 2003.

[23] J. Russo, I. Russo, *Molecular Basis of Breast Cancer: Prevention and Treatment.* 1st ed., Springer-Verlag, Heidelberg, 2004.

[24] M. Lagios, *Pathology of in-Situ Breast Cancer in Breast Cancer.* Churchill Livingstone, New York, 1999, 79-88.

[25] R. Katz-Brull, R. Margalit, P. Bendel, H. Degani, *Magn. Reson. Mater. Phys., Biol., Med.* 6 (1998) 44-52.

[26] M. Lowry, D. Manton, L. Turnbull, S. Blackband, A. Horsman, *Magn. Reson. Mater. Phys., Biol. Med.* 2 (1994) 483-485.

[27] T. E. Merchant, P. Meneses, L. W. Gierke, W. Den Otter, T. Glonek, *Br. J. Cancer* 63 (1991) 693-698.

[28] C. Mountford, S. Doran, C. Lean, P. Russell, *Chem. Rev.* 104 (2004) 3677-3704.

[29] C. Mountford, P. Russell, P. Malycha, L. Gluch, R. Somorjai, C. Lean, B. Barrraclough, D. Gillett, U. Himmelreich, B. Dolenko, E. Nikulin, C. Smith, *Br. J. Surgery* 88 (2001) 1234-1240.

[30] C. Mountford, D. Gillett, P. Russell, C. Lean, S. Doran, W. Mackinnon, J. Barr, P. Malycha, B. Barraclough, M. Bilous, *Radiology* 204 (1997) 661-666.

[31] C. Mountford, P. Russell, W. Mackinnon, E. Delikatny, *Textbook of*

Magnetic Resonance Spectroscopy in Biology and Medicine. Pergamon Press, New York, 1992.

[32] C. E. Mountford, P. Russell *Process and Device for Detecting Chemical States of Living Animal or Human Tissues by Magnetic Resonance.* US Patent: 9201946, 1992.

[33] K. K. M. Tam, R. S. Armstrong, E. A. Carter, C. E. Mountford, U. Himmelreich *Raman Spectroscopic Study of Lipid Content in Diseased Human Breast Tissues*, 4th ACOVS, Queensland University of Technology, 2001, 106-107.

[34] C. Lean, R. Somorjai, I. Smith, P. Russell, C. Mountford, *Ann. Rev. NMR. S.* 48 (2002) 71-111.

[35] C. Mountford, P. Malycha, C. Lean, S. Doran, B. Tomanek, N. Clarke, P. Stanwell, P. Russell, *Eur. Radiol.* 13 (2003) D50.

[36] C. Mountford, P. Malycha, C. Lean, R. Somorjai, B. Tomanek, L. Gluch, P. Russell, D. Gillett, S. Doran, U. Himmelreich, P. Stanwell, *Eur. Radiol.* 13 (2003) D47-49.

[37] P. Stanwell, L. Gluch, D. Clark, C. Lean, B. Giuffre, P. Malycha, B. Tomanek, C. Mountford, *Eur. Radiol.* Online Edition September (2004).

[38] C. Lean, S. Doran, R. Somorjai, P. Malycha, D. Clarke, U. Himmelreich, R. Bourne, A. Dolenikulin, C. Mountford, *Tech. Cancer Res. Treat.* Dec (2004) 551-556.

[39] K. K. M. Tam. *Spectroscopic Diagnosis of Human Female Breast Diseases*, B.Sc. Hons. Thesis, The University of Sydney, 2001.

[40] P. Russell, C. Lean, L. Delbridge, G. May, S. Dowd, C. Mountford, *Am. J. Med.* (1994) 383-388.

[41] E. J. Delikatny, S. K. Roman, R. Hancock, T. M. Jeitner, C. M. Lander, D. C. Rideout, C. E. Mountford, *Int. J. Cancer* 67 (1996) 72-79.

[42] B. R. Wood, B. Tait, D. McNaughton, *Biochim. Biophys. Acta* 1539 (2001) 58-70.

[43] H. Mantsch, M. Jackson, *Pathology by Infrared and Raman Spectroscopy in Handbook of Vibrational Spectroscopy.* Johy Wiley & Sons, Chichester, 2002, Vol. 5, 3227-3245.

[44] Kevley Technologies, http://www.kevley.com 2005.

[45] H. Fabian, W. Mäntele, *Infrared Spectroscopy of Proteins in Handbook of Vibrational Spectroscopy.* 1st ed., John Wiley & Sons, Chichester, 2002, Vol. 5, 3399-3425.

[46] H. Fabian, P. Lasch, M. Boese, W. Haensch, *Biospectroscopy* 67 (2002) 354-357.

[47] H. Fabian, P. Lasch, M. Boese, W. Haensch, *J. Mol. Struct.* 661-662 (2003) 411-417.

[48] K. Bambery, B. Wood, M. A. Quinn, D. McNaughton, *Aust. J. Chem.* 57 (2004) 1139-1143.

[49] I. Levin, R. Bhargava, *Annu. Rev. Phys. Chem.* 56 (2005) 429-474.

[50] N. Colthup, L. Daly, S. Wiberley, *Introduction to Infrared and Raman Spectroscopy.* 2nd ed., Academic Press, London, 1975.

[51] D. Long, *Raman Spectroscopy.* McGraw-Hill, London, 1977.

[52] B. Schrader, *Infrared and Raman Spectroscopy: Methods and Applications.* 1st ed., VCH Publishers, New York, 1995.

[53] K. Nakamoto, *Infrared and Raman Spectra of Inorganic and Coordination Compounds.* 5th ed., Wiley, New York, 1997.

[54] G. Socrates, *Infrared and Raman Characteristic Group Frequencies: Tables and Charts.* 3rd ed., John Wiley & Sons Ltd., Chichester, 2001.

[55] P. Atkins, *Physical Chemistry.* 6th ed., Oxford University Press, Melbourne, 1998.

[56] K. K. M. Tam, E. A. Carter, P. A. Lay, H. H., R. S. Armstrong, C. E. Mountford, S. Dowd, P. Russell, Y. C. Lee, Y. L. Chen *Characterisation of Ductal Carcinoma in-Situ Breast Cancer Using IR Microspectroscopy,* 6th ACOVS, The University of Sydney, 2005, 39-40.

[57] M. Jackson, K. Kim, J. Tetteh, J. R. Mansfield, B. Dolenko, R. L. Somorjai, F. W. Orr, P. H. Watson, H. H. Mantsch, *Proc. SPIE-Int. Soc. Opt. Eng.* 3257 (1998) 24-34.

[58] T. Gao, J. Feng, Y. Ci, *J. Eur. Soc. Anal. Cell. Path.* 18 (1999) 87-93.

[59] T. Gao, Y. Ci, *Anal. Commun.* 36 (1999) 341-343.

[60] H. Fabian, R. Wessel, M. Jackson, A. Schwartz, P. Lasch, I. Fichtner, H. H.

Mantsch, D. Naumann, *Proc. SPIE-Int. Soc. Opt. Eng.* 3257 (1998) 13-23.

[61] Y. Ci, T. Gao, J. Feng, Z. Guo, X. Kan, J. Dong, *J. Chin. Sci. Bull.* 44 (1999) 804-808.

[62] Y. Ci, T. Gao, J. Feng, Z. Guo, *Appl. Spectrosc.* 53 (1999) 312-315.

[63] R. K. Dukor, M. N. Liebman, B. L. Johnson, *Cell. Mol. Biol.* 44 (1998) 211-217.

[64] B. Rigas, P. T. T. Wong, *Cancer Res.* 52 (1992) 84-88.

[65] S. Singer, K. Souza, W. Thilly, *Cancer Res.* 55 (1995) 5140-5145.

[66] K. L. Worden, Y. Song, X. Jiang, A. Constantinescu, R. P. Mason, H. Liu, *Proc. SPIE-Int. Soc. Opt. Eng.* 3597 (1999) 601-610.

[67] M. Jackson, L. Choo, P. H. Watson, W. C. Halliday, H. H. Mantsch, *Biochim. Biophys. Acta* 1270 (1995) 1-6.

[68] L. M. McIntosh, M. Jackson, H. H. Mantsch, M. F. Stranc, D. Pilavdzic, A. N. Crowson, *J. Invest. Dermatol.* 112 (1999) 951-956.

[69] D. C. Malins, N. L. Polissar, K. Nishikida, E. H. Holmes, H. S. Gardner, S. J. Gunselman, *Cancer* 75 (1995) 503-517.

[70] H. Huo, X. Hu, H.-W. Guan, X.-F. Wang, X. Che, W.-D. Huang, *Gaodeng Xuexiao Huaxue Xuebao* 21 (2000) 1244-1247.

[71] S. Shen, B. Liu, X. Ma, Z. Song, Q. Li, *Guangpuxue Yu Guangpu Fenxi* 20 (2000) 28-30.

[72] R. Eckel, H. Huo, H.-W. Guan, X. Hu, X. Che, W.-D. Huang, *Vib. Spectrosc.* 27 (2001) 165-173.

[73] M. Meurens, J. Wallon, J. Tong, H. Noeel, J. Haot, *Vib. Spectrosc.* 10 (1996) 341-346.

[74] B. Wood, L. Chiriboga, H. Yee, M. A. Quinn, D. McNaughton, M. Diem, *Gynecol. Oncol.* 93 (2004) 59-68.

[75] B. R. Wood, M. A. Quinn, B. Tait, M. Ashdown, T. Hislop, M. Romeo, D. McNaughton, *Biospectroscopy* 4 (1998) 75-91.

[76] N. I. Afanasyeva, V. Artjushenko, S. Kolyakov, V. Letokhov, A. Lerman, G.

Nikolaev, V. Sokolov, *Proc. SPIE-Int. Soc. Opt. Eng.* 2627 (1995) 228-230.

[77] G. Budinova, K. Volka, *Chem. Listy* 90 (1996) 433-445.

[78] P. Lasch, W. Haensch, D. Naumann, M. Diem, *Biochem. Biophys. Acta* 1688 (2004) 176-186.

[79] M. Jackson, H. H. Mantsch, *Pathology by Infrared and Raman Spectroscopy in Handbook of Vibrational Spectroscopy.* John Wiley & Sons, Chichester, 2002, Vol. 5, 3227-3245.

[80] H. Fabian, M. Jackson, L. Murphy, P. H. Watson, I. Fichtner, H. H. Mantsch, *Biospectroscopy* 1 (1995) 37-45.

[81] M. Diem, J. Bargonetti, T. Gopen, S. Boydston-White *Method and Apparatus for Characterization of Biological Entities Such as Cells Using Ir Spectra.* US Patent: 0043775, 2000.

[82] P. Lasch, D. Naumann, *Biochim. Biophys. Acta* 1758 (2006) 814-829.

[83] L. Miller, R. Smith, *Vib. Spectrosc.* 38 (2005) 237-240.

[84] L. Miller, *Microsc. Microanal.* 10 (Suppl. 2) (2004) 1320-1321.

[85] P. Dumas, *Microsc. Microanal.* 10 (Suppl. 2) (2004) 1322-1323.

[86] L. Miller, *J. Biol. Phys.* 29 (2003) 219-230.

[87] S. O. Miller, G. P. Ewing, C. Howard, H. Tachikawa, S. A. Bigler, W. H. Barber, M. Angel, D. O. McDaniel, *Biomed. Sci. Instrum.* 39 (2003) 24-29.

[88] L. Miller, P. Dumas, *Biochem. Biophys. Acta* 1758 (2006) 846-857.

[89] M. Jackson, J. R. Mansfield, B. Dolenko, R. L. Somorjai, H. H. Mantsch, *Cancer Detect. Prev.* 23 (1999) 245-253.

[90] C. P. Schultz, K.-Z. Liu, E. A. Salamon, K. T. Riese, H. H. Mantsch, *J. Mol. Struct.* 480-481 (1999) 369-377.

[91] C. P. Schultz, H. H. Mantsch, *Cell. Mol. Biol.* 44 (1998) 203-210.

[92] H. Mantsch, M. Jackson, *J. Mol. Struct.* 347 (1995) 187-206.

[93] C. P. Schultz, H. H. Eysel, H. H. Mantsch, M. Jackson, *J. Phys. Chem.* 100 (1996) 6845-6848.

[94] M. G. Sowa, J. J. Payette, H. H. Mantsch, *J. Surg. Res.* 86 (1999) 62-69.

[95] T. Gao, Y. Ci, *Internet J. Vib. Spectrosc.* 3 (1999) No pp given.

[96] B. Stuart, W. George, P. McIntyre, *Modern Infrared Spectroscopy: Analytical Chemistry by Open Learning.* John Wiley & Sons, Chichester, 1996.

[97] H. Dehghani, B. Brooksby, B. Pogue, K. Paulsen, *Appl. Opt.* 44 (2005) 1870-1878.

[98] Y. Xu, Y. Zhao, Z. Xu, Y. Ren, Y. Liu, Y. Zhang, X. Zhou, J. Shi, D. Xu, J. Wu, *Guangpuxue Yu Guangpu Fenxi* 25 (2005) 1775-1778.

[99] G. Yu, S. Lu, J. Xu, C. Zhang, C. Zhang, *Guangzi Xuebao* 34 (2005) 390-394.

[100] G. Yu, S. Lu, J. Xu, C. Zhang, G. Zhang, *Nankai Daxue Xuebao, Ziran Kexueban* 38 (2005) 53-56.

[101] G. Yu, J. Xu, Y. Niu, C. Zhang, C. Zhang, *Proc. SPIE-Int. Soc. Opt. Eng.* 5630 (2005) 796-801.

[102] J. Mansfield, L. McIntosh, A. Crowson, H. Mantsch, M. Jackson, *Appl. Spectrosc.* 53 (1999) 1323-1330.

[103] M. Diem, L. Chiriboga, H. Yee, *Biopolymers* 57 (2000) 282-290.

[104] M. Jackson, B. Ramjiawan, M. Hewko, H. H. Mantsch, *Cell. Mol. Biol.* 44 (1998) 89-98.

[105] P. Lasch, D. Naumann, *Cell. Mol. Biol.* 44 (1998) 189-202.

[106] J. R. Mansfield, M. G. Sowa, G. B. Scarth, R. L. Somorjai, H. H. Mantsch, *Anal. Chem.* 69 (1997) 3370-3374.

[107] Cytospec, *Software: Cytospec, Version 1.3*; 2006.

[108] M. Diem, M. Romeo, *Vib. Spectrosc.* 38 (2005) 115-119.

[109] H. Fabian, N. Thi, M. Eiden, P. Lasch, J. Schmitt, D. Naumann, *Biochem. Biophys. Acta* 1758 (2006) 874-882.

[110] L. Zhang, G. Small, A. Haka, L. Kidder, E. Lewis, *Appl. Spectrosc.* 57 (2003) 14-22.

[111] P. Dumas, L. Miller, *Vib. Spectrosc.* 32 (2003) 3-21.

[112] M. Tobin, J., M. Chesters, A., J. Chalmers, M., F. Rutten, J. M., S. Fisher, E., I. M. Symonds, A. Hitchcock, R. Allibone, S. Dias-Gunasekara, *Faraday Discuss.* 126 (2004) 27-39.

[113] E. Gazi, N. P. Lockyer, J. C. Vickerman, P. Gardner, J. Dwyer, C. A. Hart, M. D. Brown, N. W. Clarke, J. Miyan, *Appl. Surf. Sci.* 231-232 (2004) 452-456.

[114] E. O. Faolain, M. B. Hunter, J. M. Byrne, P. Kelehan, H. J. Byrne, F. M. Lyng, *Proc. SPIE-Int. Soc. Opt. Eng.* 5826 (2005) 25-36.

[115] M. Diem, M. Romeo, C. Matthaus, M. Miljkovic, L. Miller, P. Lasch, *Infrared Phys. Techn.* 45 (2004) 331-338.

[116] D. Wetzel, G. Post, R. Lodder, *Vib. Spectrosc.* 38 (2005) 53-59.

[117] E. Gazi, J. Dwyer, N. P. Lockyer, J. Miyan, P. Gardner, C. A. Hart, M. D. Brown, N. W. Clarke, *Vib. Spectrosc.* 38 (2005) 193-201.

[118] C. Liu, Y. Zhang, X. Yan, X. Zhang, C. Li, W. Yang, D. Shi, *J. Lumin.* 119-120 (2006) 132-136.

[119] R. N. Baker, J. Smith, K. D. Rogers, N. Stone, *Proc. SPIE-Int. Soc. Opt. Eng.* 6093 (2006) 1-11.

[120] D. Pappas, B. W. Smith, J. D. Winefordner, *Talanta* 51 (2000) 131-144.

[121] E. E. Lawson, H. G. M. Edwards, B. W. Barry, A. C. Williams, *J. Drug Targeting* 5 (1998) 343-351.

[122] A. Haka, K. Shafer-Peltier, M. Fitzmaurice, J. Crowe, R. Dasari, M. Feld, *Cancer Res.* 62 (2002) 5375-5380.

[123] J. Kneipp, T. C. Bakker Schut, M. Kliffen, M. Menke-Pluijmers, G. J. Puppels, *Vib. Spectrosc.* 32 (2003) 67-74.

[124] Y. Yazdi, N. Ramanujam, R. Lotan, M. F. Mitchell, W. Hittelman, R. Richards-Kortum, *Appl. Spectrosc.* 53 (1999) 82-85.

[125] J. Tong, M. Meurens, H. Noel, *Near Infrared Spectrosc.: Future Waves, Proc. Int. Conf. Near Infrared Spectrosc., 7th* (1996) 334-336.

[126] K. Shafer-Peltier, R. Manoharan, R. Dasari, M. S. Feld, G. R. Harrison, J. Myles, J. Crowe, M. Fitzmaurice, *Book of Abstracts, 216th ACS National Meeting, Boston, August 23-27* (1998) HYS-344.

[127] B. Schrader, S. Keller, T. Loechte, S. Fendel, D. S. Moore, A. Simon, J. Sawatzki, *J. Mol. Struct.* 348 (1995) 293-296.

[128] B. Schrader, B. Dippel, S. Fendel, S. Keller, T. Lochte, M. Riedl, R. Schulte, E. Tatsch, *J. Mol. Struct.* 408–409 (1997) 23-31.

[129] M. D. Schaeberle, V. F. Kalasinsky, J. L. Luke, E. N. Lewis, I. W. Levin, P. J. Treado, *Anal. Chem.* 68 (1996) 1829-1833.

[130] D. C. B. Redd, Z. C. Feng, K. T. Yue, T. S. Gansler, *Appl. Spectrosc.* 47 (1993) 787-791.

[131] R. P. Rava, J. J. Baraga, M. S. Feld *Systems and Methods of Molecular Spectroscopy to Provide for the Diagnosis of Tissue*. Patent: 9215008, 1992.

[132] J. D. Pasteris, B. Wopenka, J. J. Freeman, V. L. Young, H. J. Brandon, *Am. Mineral.* 84 (1999) 997-1008.

[133] R. Manoharan, Y. Wang, M. S. Feld, *Spectrochim. Acta, Part A* 52A (1996) 215-249.

[134] R. Manoharan, K. Shafer-Peltier, L. Perelman, J. Wu, K. Chen, G. Deinum, M. Fitzmaurice, J. Myles, J. Crowe, R. Dasari, M. Feld, *Photochem. Photobiol.* 67 (1998) 15-22.

[135] L. H. Kidder, V. F. Kalasinsky, J. L. Luke, I. W. Levin, E. N. Lewis, *Nat. Med. (N. Y.)* 3 (1997) 235-237.

[136] C. J. Frank, D. C. B. Redd, T. S. Gansler, R. L. McCreery, *Anal. Chem.* 66 (1994) 319-326.

[137] C. J. Frank, R. L. McCreery, D. C. B. Redd, T. S. Gansler, *Appl. Spectrosc.* 47 (1993) 387-390.

[138] C. J. Frank, R. L. McCreery, D. C. B. Redd, *Anal. Chem.* 67 (1995) 777-783.

[139] M. S. Feld, R. Manoharan, J. Salenius, J. Orenstein-Carndona, T. J. Romer, J. F. Brennan, III, R. Dasari, Y. Wang, *Proc. SPIE-Int. Soc. Opt. Eng.* 2388 (1995) 99-104.

[140] K. Shafer-Peltier, A. Haka, M. Fitzmaurice, J. Crowe, J. Myles, R. Dasari, M. Feld, *J. Raman Spectrosc.* 33 (2002) 552-563.

[141] R. Alfano, C. Liu, W. Sha, H. Zhu, D. Akins, J. Cleary, R. Prudente, E. Clemer, *Lasers Life Sci.* 4 (1991) 23-28.

[142] E. Hanlon, R. Manoharan, T. Koo, K. Shafer, J. Motz, M. Fitzmaurice, J. Kramer, I. Itzkan, R. Dasari, M. Feld, *Phys. Med. Biol.* (2000) R1-R59.

[143] S. Fendel, B. Schrader, *Fresenius' J. Anal. Chem.* 360 (1998) 609-613.

[144] G. J. Puppels, J. Greve, *in Biomedical Applications of Spectroscopy.* John Wiley & Sons Ltd., 1996, Vol. 25, 1-47.

[145] J. F. Brennan, III, Y. Wang, R. R. Dasari, M. S. Feld, *Appl. Spectrosc.* 51 (1997) 201-208.

[146] M. Chowdary, K. Kumar, J. Kurien, S. Mathew, C. Krishna, *Biopolymers* 83 (2006) 556-569.

[147] K. Shafer-Peltier, A. Haka, J. Motz, M. Fitzmaurice, R. Dasari, M. Feld, *J. Cell. Biochem.* 39 (2002) 125-137.

[148] Matteo Matteucci, A Tutorial on Clustering Algorithms. http://www.elet.polimi.it/upload/matteucc/Clustering/tutorial_html/index.html 2006.

[149] T. Næs, T. Isaksson, T. Fearn, T. Davies, *A User-Friendly Guide to Multivariate Calibration and Classification.* NIR Publications, Chicester, 2002.

[150] M. Otto, *Chemometrics: Statistics and Computer Application in Analytical Chemistry.* 1st ed., Wiley-VCH, Weinheim, 1999.

[151] R. Brereton, *Chemometrics: Data Analysis for the Laboratory and Chemical Plant.* 1st ed., John Wiley & Sons, Chichester, 2003.

[152] M. Adams, *Chemometrics in Analytical Spectroscopy.* 2nd ed., Cambridge: Royal Society of Chemistry, Cambridge, 2004.

[153] J. B. McQueen *Some Methods for Classification and Analysis of Multivariate Observations,* 5th Berkeley Symposium on Mathematical Statistics and Probability, L. LeCam, J. Neymann, Eds. University of California Press, Berkeley, 1967, 281-297.

[154] J. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms.* Plenum Press, New York, 1981.

2. MATERIALS AND METHODS

2.1 BREAST SPECIMENS

One hundred breast specimens were obtained from routine surgical biopsies (both fine-needle aspirated and open-core excisional biopsies) with the ethics clearance approvals (The Sydney Adventist Hospital: approval number: 10/00, Sydney, NSW; and Royal Adelaide Hospital: Surgeon Dr. Peter Malycha, approval numbers: 940424b and 970308c, Adelaide, SA, Australia). The breast specimens were placed in polypropylene vials containing 2 mL of deuterated phosphate buffer saline solution (PBS/D_2O, pH 7.0, ICN Biomedicals Inc., Ohio, USA), snapped frozen in liquid nitrogen and subsequently stored at –70 °C until they were required for histopathology, magnetic resonance (parallel study) and vibrational spectroscopic analysis.

Samples were stored and examined typically within six months from the date of acquisition. Twenty of the 100 samples were previously analysed using magnetic resonance spectroscopy (MRS) and were over two years old at the time of the vibrational spectroscopic analyses.

2.2 TISSUE PREPARATION AND HISTOPATHOLOGY

Tissue harvested for histopathology, infrared and Raman spectroscopic analyses was sectioned by one of two methods, cryosectioning or formalin-fixed paraffin-embedded (FFPE) sectioning, see Figure 2.1. Fine needle aspirated (FNA) breast samples were collected primarily for a parallel study using MRS at IMRR. The tissue sectioning and staining were performed in the Department of Histopathology, The University of Sydney and were examined by an expert pathologist Prof. Peter Russell (PR), The Department of Anatomical Histopathology, Royal Prince Alfred Hospital (RPAH), Sydney.

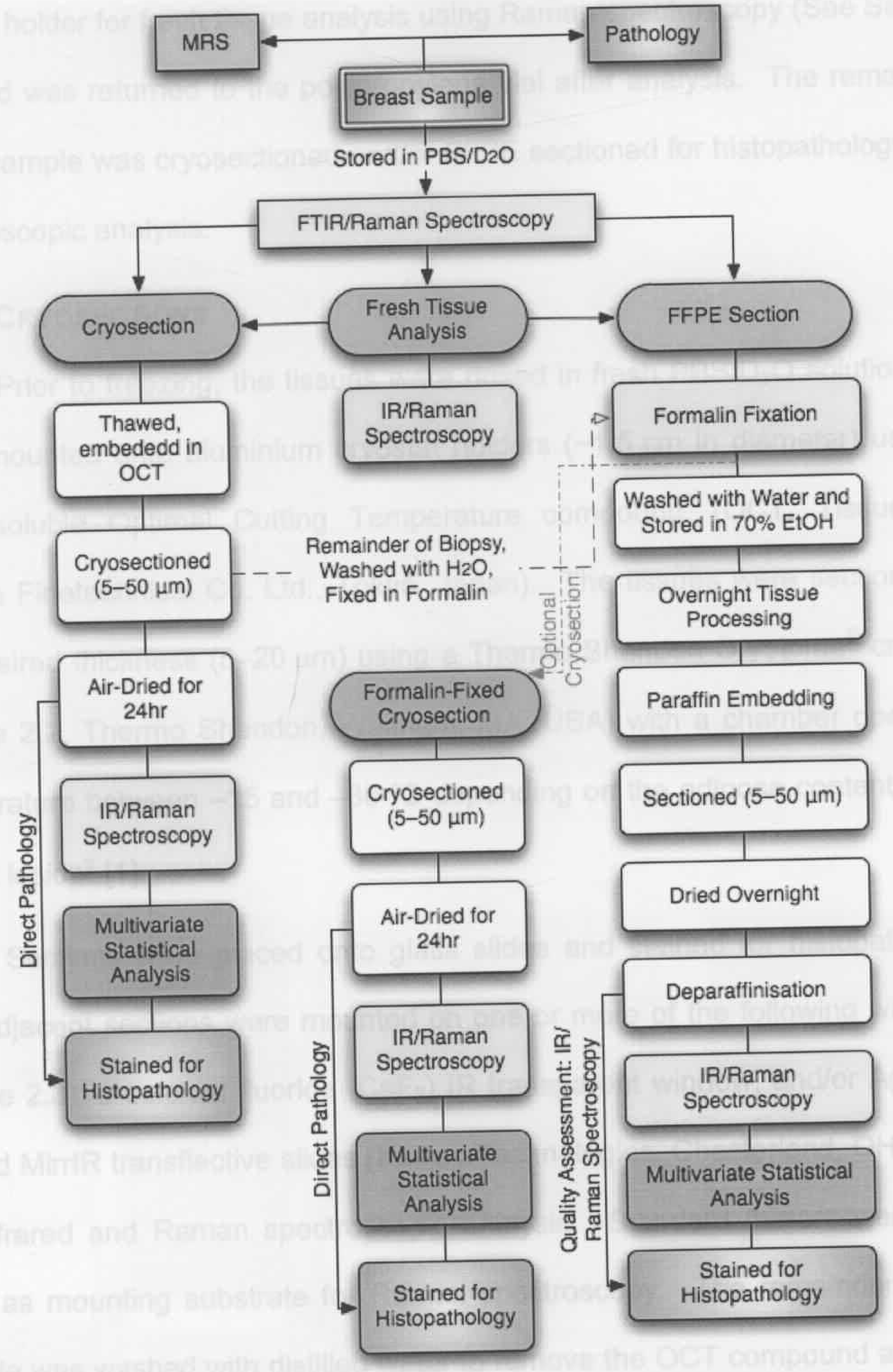


Figure 2.1. Schematic workflow of sample preparation and experimentation.

2.2.1 FRESH TISSUES

Fresh biopsy tissues were stored in PBS/D_2O solution at –70 °C with no further chemical treatment (minimum H_2O interference for MRS experiments). One third of each sample (c.a. 1 mm^3 in volume) was placed into an aluminium sample holder for fresh tissue analysis using Raman spectroscopy (See Section 2.4) and was returned to the polypropylene vial after analysis. The remainder of the sample was cryosectioned and/or FFPE sectioned for histopathology and spectroscopic analysis.

2.2.2 CRYOSECTIONS

Prior to freezing, the tissues were rinsed in fresh PBS/D_2O solution and were mounted onto aluminium cryostat holders (~1.5 cm in diameter) using a water-soluble Optimal Cutting Temperature compound* (OCT, Tissue-Tek, Sakura Finetechnical Co. Ltd., Tokyo, Japan). The tissues were sectioned to the desired thickness (5–20 μm) using a Thermo Shandon Cryotome® cryostat (Figure 2.2, Thermo Shandon, Waltham, MA, USA) with a chamber operating temperature between –15 and –30 °C depending on the adipose content in the breast lesion† [1].

Sections were placed onto glass slides and stained for histopathology and adjacent sections were mounted on one or more of the following windows (Figure 2.3): a calcium fluoride (CaF_2) IR transparent window, and/or Ag/SnO_2 coated MirrIR transflective slides (Kevley Technologies, Chesterland, OH, USA) for infrared and Raman spectroscopic analysis. Standard mirrors were also used as mounting substrate for Raman spectroscopy. The remainder of the sample was washed with distilled water to remove the OCT compound and was placed into a new polypropylene vial containing 10% formalin solution overnight for FFPE sectioning [1]. Some cryosections were also prepared on biopsies after formalin-fixation for the sample optimisation study.

---

[*] OCT ingredients: 10.24% w/w polyvinyl alcohol, 4.26% w/w polyethylene glycol and 85.5% w/w non-reactive ingredients.

[†] Recommended operating temperature for tissue with high adipose content: –20 to –30 °C and for low adipose content: –15 to –20 °C.

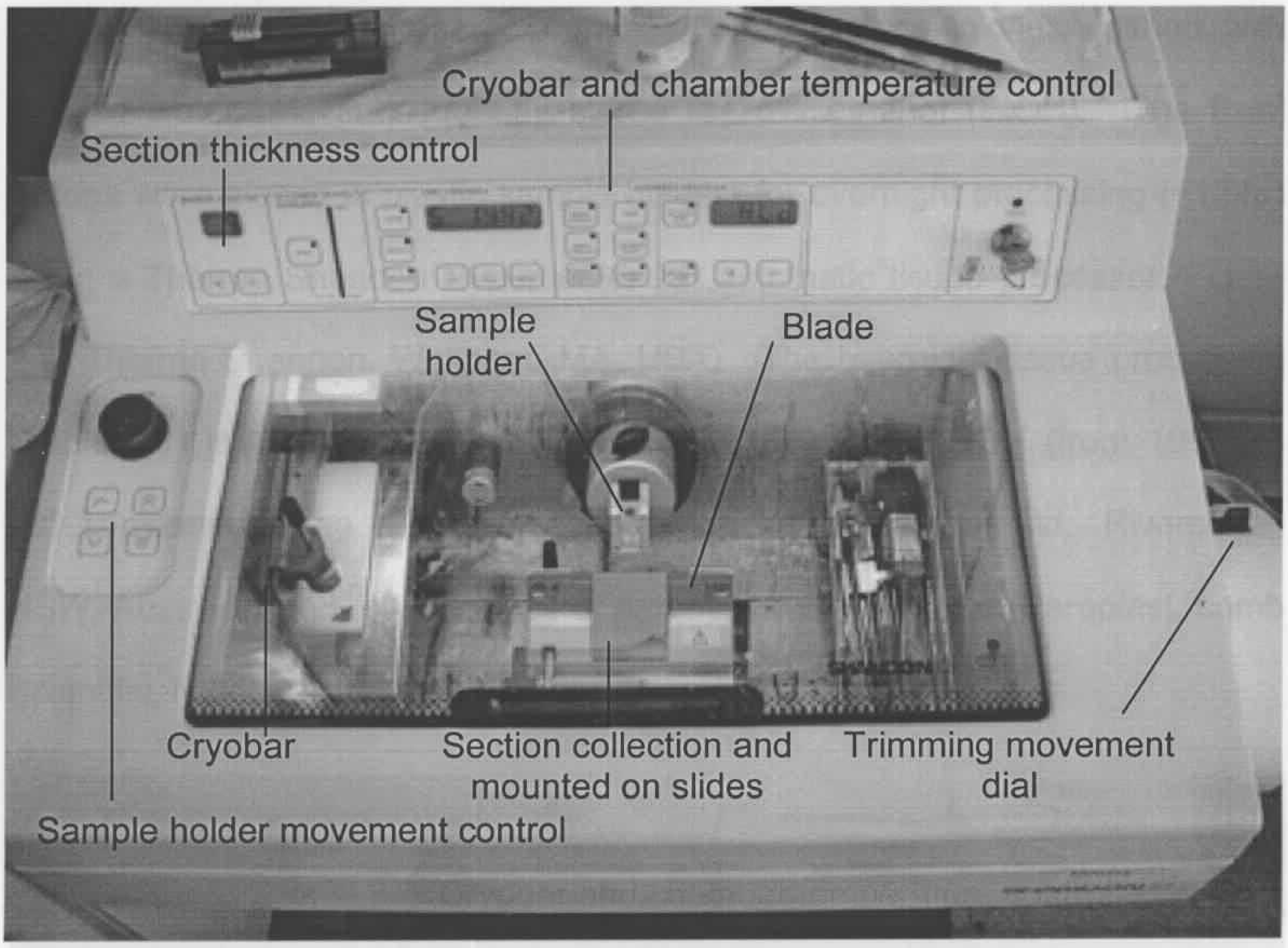


Figure 2.2. Thermo Shandon Cryotome® used for tissue cryosectioning.

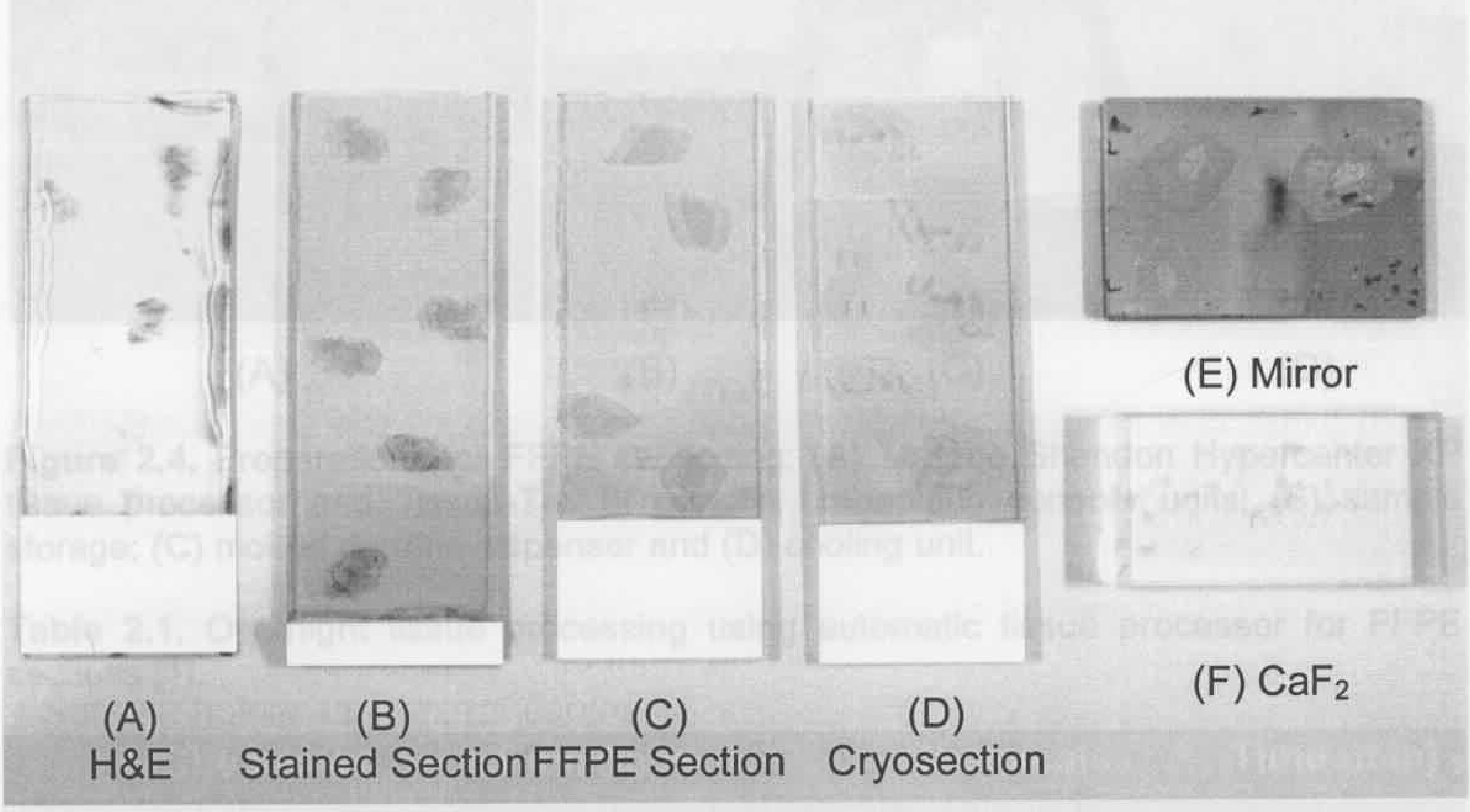


Figure 2.3. Tissue mounting materials: (A) glass containing H&E stained sections; transflective slides containing (B) Nuclei-stained sections; (C) FFPE sections; (D) cryosections; (E) cryosections on mirror; and (F) cryosections on CaF_2 window.

2.2.3 FORMALIN-FIXED PARAFFIN-EMBEDDED SECTIONS

Prior to paraffin embedding, the biopsies were fixed in a 10% formalin solution (4% formaldehyde, 10% v/v, Sigma-Aldrich Pty. Ltd., USA) for a minimum of 24 hr[‡] to ensure complete protein cross-linking, washed with distilled water and subsequently stored in 70% ethanol (EtOH). The fixed tissues were placed in plastic sample holders for overnight processing (~13 hr) using a Thermo Shandon Hypercenter XP automatic tissue processor (Figure 2.4, Thermo Shandon, Waltham, MA, USA). The overnight tissue processing (Table 2.1) involved tissue dehydration through graded EtOH (from 100% to 70%), clearing using two changes of xylene (Fronine Pty. Ltd., Riverstone, NSW, Australia) and four changes of paraffin wax infiltrations (Paraplast, Lomb Scientific, NSW, Australia) [1].

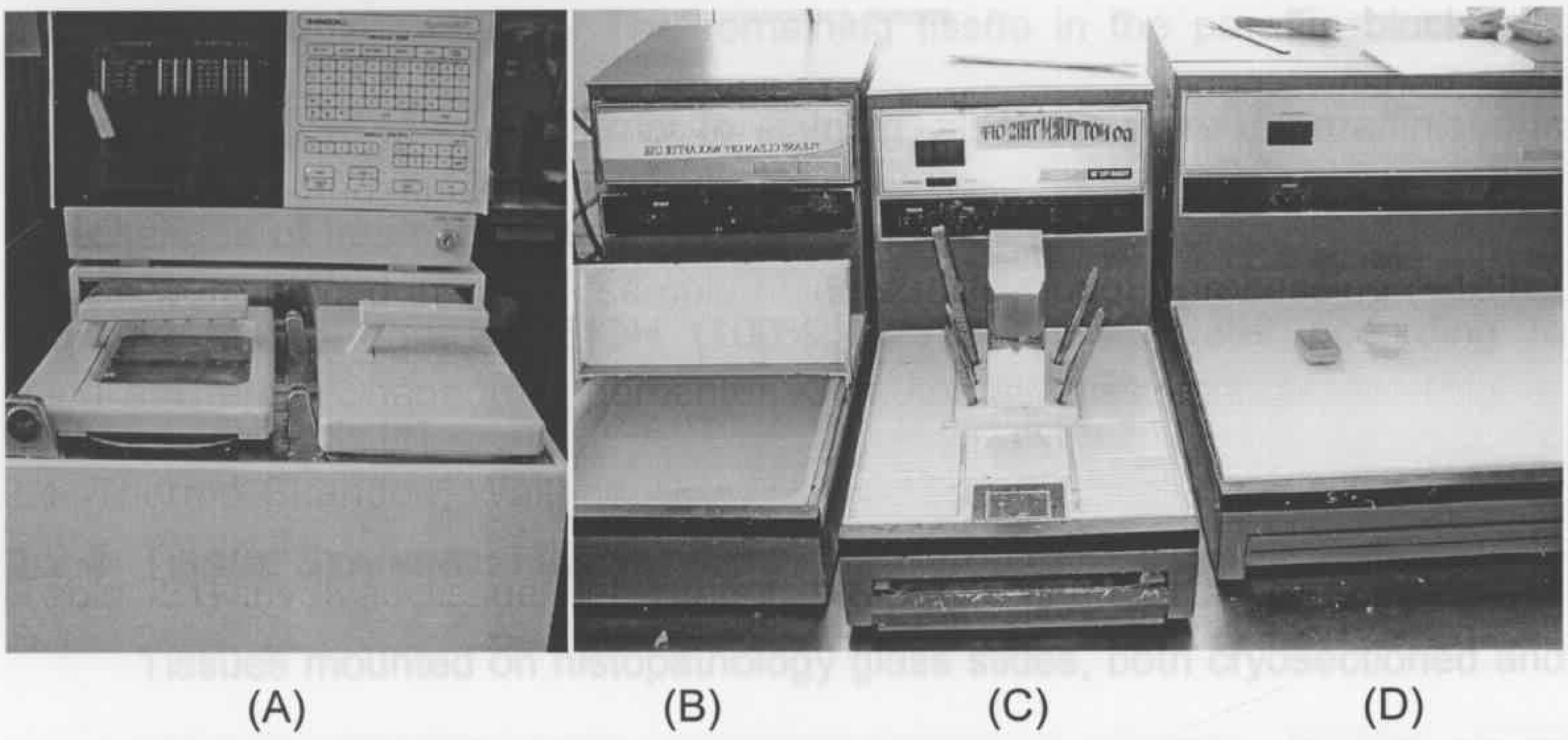


Figure 2.4. Preparations for FFPE sectioning: (A) Thermo Shandon Hypercenter XP tissue processor and Tissue-Tek III paraffin embedding console units: (B) sample storage; (C) molten paraffin dispenser and (D) cooling unit.

Table 2.1. Overnight tissue processing using automatic tissue processor for FFPE sections [1].

	Chemicals	Time (min)		Chemicals	Time (min)
1	EtOH 50%	40	8	Xylene I	60
2	EtOH 70%	40	9	Xylene II	60
3	EtOH 80%	40	10	Paraffin I	90
4	EtOH 95%	60	11	Paraffin II	60
5	EtOH 95%	60	12	Paraffin III	60
6	EtOH 100%	60	13	Paraffin IV	90
7	EtOH 100%	60			

[‡] A large biopsy requires a longer fixation time prior to FFPE sectioning.

The tissues were embedded and oriented in aluminium sample holders, waxed with molten paraffin at 62 °C (Figure 2.4C), and cooled on ice for at least 30 min (Figure 2.4D). The paraffin block was then placed onto a Thermo Shandon Finesse 325 microtome (Thermo Shandon, Waltham, MA, USA) and sectioned to the desired thicknesses of between 5 and 50 μm. The sections were placed into a water bath at 42 °C and were mounted on histopathology glass slides (5 to 10 μm). Adjacent sections were placed onto CaF_2, infrared transflective slides and/or mirrors for infrared and Raman spectroscopic analysis. All sections were dried in an oven overnight at 45 °C before staining and spectroscopic analysis. The remaining tissue in the paraffin block was sealed and stored at –4 °C. Prior to staining, all slides were deparaffinised in two changes of fresh Histopure or xylene solutions (10 min each) and then re-hydrated through graded EtOH (100% to 70%) then water according to standard protocols [1].

2.2.4 TISSUE STAINING: HAEMATOXYLIN AND EOSIN

Tissues mounted on histopathology glass slides, both cryosectioned and FFPE sectioned, were stained using a standard haematoxylin and eosin (H&E) procedure [1]. A detailed staining procedure is summarised in Table 2.2, and subtle differences used in sample preparations are as followed.

The FFPE sections were first deparaffinised in two changes of xylene (Step 1) and rehydrated through treatments with graded EtOH (100% to 70%, Step 2), while cryosections were rehydrated only in 70% EtOH. The use of ethanol in the H&E process removes adipose/lipid content within the sections (both cryosections and FFPE sections). Sectioned tissues mounted on IR transflective slides were also stained after IR spectroscopic analysis.

Table 2.2. Haematoxylin and eosin tissue staining procedures [1].

	H&E Protocol	Time
1	Deparaffinise sections in two changes of xylene or histoclear solution (Fronine Pty. Ltd., Riverstone, NSW, Australia for paraffin-sections only)	10 min each
2	Dehydrate sections through graded EtOH (100%, 100%, 95%, 95%, 70%) and tap water	2 min each
3	Stain in Harris's hematoxylin (Fronine Pty. Ltd., Riverstone, NSW, Australia)	3–4 min
4	Wash in tap water	1 min
5	Differentiate in acid alcohol (70% EtOH, 4 L and 10 M HCl, 10 mL) to remove excess stain	3–4 dips
6	Wash in water	1 min
7	Stain in Scott's blueing solution (Sigma-Aldrich, USA)	30 sec
8	Examine under microscope to assess whether the nuclei are clearly stained and the cytoplasm is unstained. If not completely differentiated, repeat steps 4–6	
9	Wash in water	2–5 min
10	Dip briefly in 70% EtOH	3–4 dips
11	Counterstain in two changes of eosin solution (Fronine Pty. Ltd., Riverstone, NSW, Australia)	30–40 sec each
12	Dehydrate through graded alcohols, 95%, 95%, 100% and 100%	3–4 dips
13	Clear in two changes of xylene or histoclear solution and cover slip using DPX solution (BDH Gurr, Crown Scientific Australia)	2 min each

2.3 FOURIER TRANSFORM INFRARED SPECTROSCOPY

2.3.1 FTIR MICROSCOPY: TISSUE SECTION ANALYSIS

Infrared spectra were recorded using a Bruker IFS66v FTIR spectrometer (Bruker GmbH, Rudolf-Plank-Straβe, Germany) equipped with a Ge-coated KBr beamsplitter with a spectral range of 4500–370 cm^{-1} (mid-infrared region, MIR) and a room temperature deuterated L-alanine triglycerine sulfate (DLaTGS) detector (Figure 2.5). Infrared spectra of sectioned tissues were obtained using a Bruker Series One Microscope with a liquid-nitrogen-cooled MCT detector attached to the IFS66v FTIR spectrometer. The microscope is equipped with a trinocular viewer, which accommodates a video camera allowing direct viewing of the sample and three objectives; one for viewing the sample (×4/0.10 NA) and the other two for both viewing and infrared measurements (×15/4.0 NA and ×36/0.50 NA). The microscope works in both transmittance and reflectance modes. A cassegrainian lens is permanently aligned and used to focus the beam on the MCT detector. A *x,y* sample stage (160 × 155 mm) was manually driven enabling viewing of specific sample areas and collection of spectra.

IR spectra of sectioned samples (5–50 μm, typically 5 μm) mounted on CaF_2 windows or Ag/SnO_2 coated IR transflective slides were collected using a ×15 objective with a fully-opened knife edge aperture of 5.0 × 5.0 mm^2. The background spectrum was collected from a clean infrared window without sample. The spectrometer bench was evacuated and the microscope sample area was purged with constant flow of nitrogen gas to minimise atmospheric CO_2 and water vapour contributions. All spectra (breast sections and biological reference samples) were collected over the MIR region with the co-addition of 256 or 512 scans at a spectral resolution of 4 cm^{-1}. All single-point experiments involved the acquisition of spectra in triplicate. At least two areas on each sample were investigated with triplicate measurements collected within each area to assess reproducibility and repeatability.

IR experimental data were collected using Bruker OPUS™ software [2] (*.0 extension, version 5.0.53, Bruker Analytik, Germany). Triplicate spectra were averaged and converted to GRAMS™ format (*.spc extension, version 7.02, Thermo Galactic, Salem USA) for spectral analysis in GRAMS™ software [3].

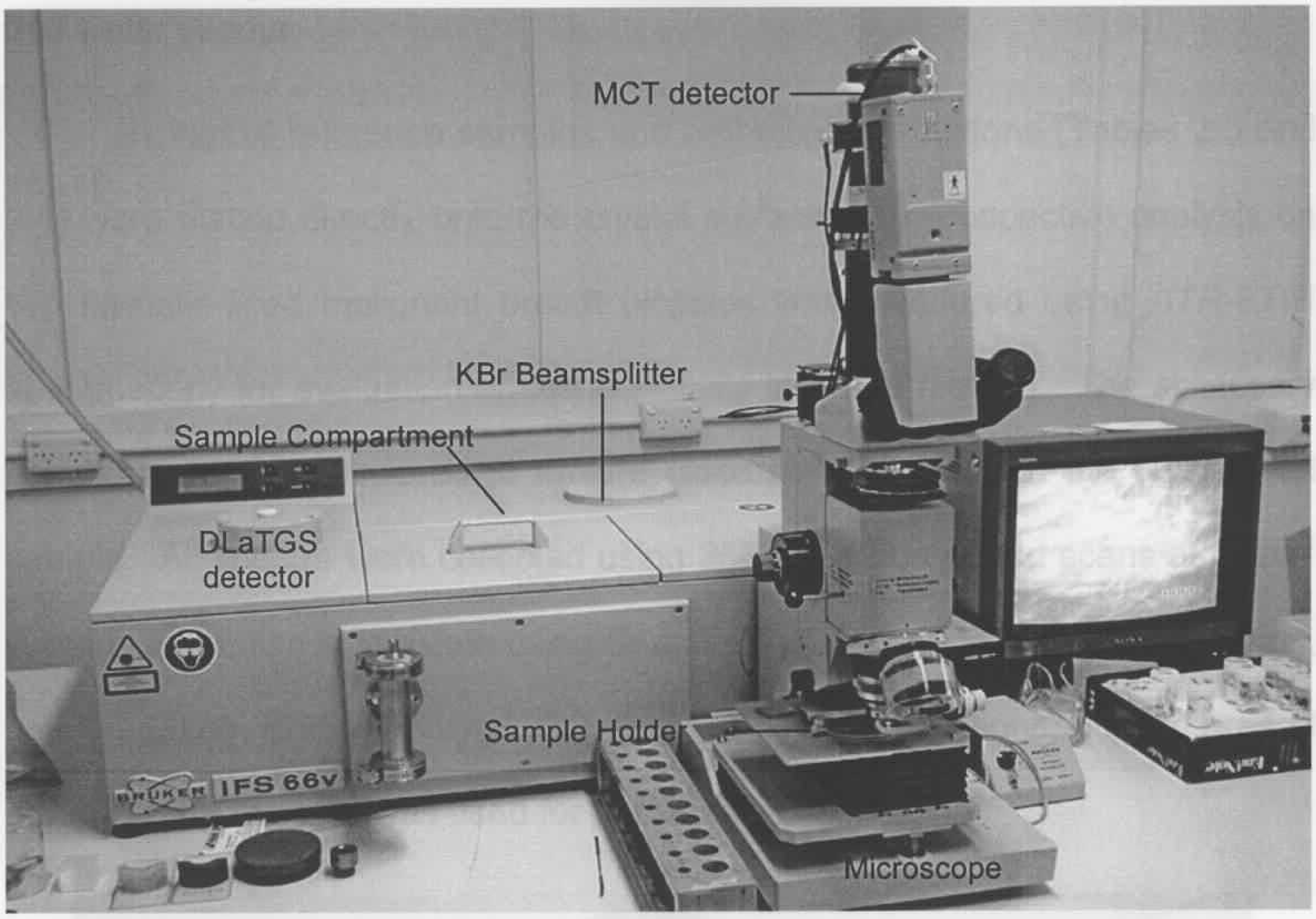


Figure 2.5. Bruker IFS66v FTIR spectrometer with Series One Microscope.

2.3.2 ATTENUATED TOTAL REFLECTION SPECTROSCOPY

Attenuated total reflection (ATR) or internal reflection spectroscopy (IRS) is a versatile (minimal sample preparation) sampling technique. The ATR accessory (Figure 2.6) used was a MIRacle™ single reflection horizontal ATR (Pike Technologies, Madison, WI, USA) equipped with a composite diamond internal reflection element (IRE) with a 2-mm sampling surface and a ZnSe focusing element. The composite diamond/ZnSe IRE has a refractive index of 2.403. A single-beam background spectrum was collected from a clean, blank IRE surface. A single-beam sample spectrum was then collected and the ratio of these spectra was used to produce a spectrum in absorbance units. After each measurement, the IRE was thoroughly cleaned with a cotton bud dipped in isopropanol and allowed to air-dry prior to the collection of a new background spectrum. The ATR accessory was purged with nitrogen prior to and during spectral acquisition to minimise atmospheric contributions from carbon dioxide and water vapour.

Biological reference samples and pathological solutions (Tables 2.3 and 2.4) were placed directly onto the crystal surface. A retrospective analysis on two formalin-fixed malignant breast biopsies was measured using ATR-FTIR spectroscopy for evaluation of tissue stability in tissue fixation. The accessory sample clamp was lowered to ensure close contact between the crystal and sample. All spectra were collected using 256 or 512 co-added scans at 4 cm^{-1} spectral resolution in triplicate using Bruker OPUS™ software [2] (*.0 extension, version 5.0.53, Bruker Analytik, Germany). For breast sections, a clean area of the mounting material was used for background collection.

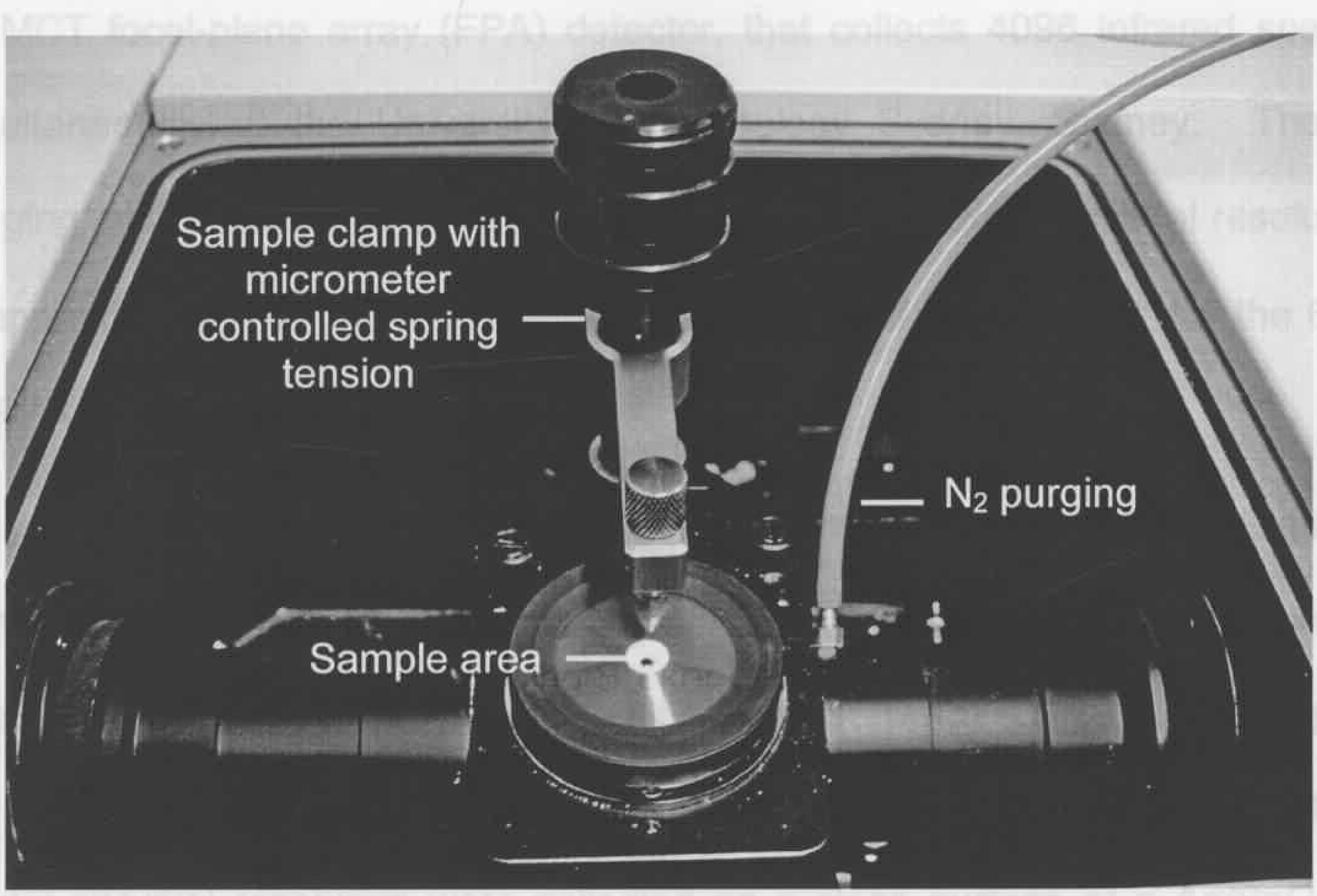


Figure 2.6. ATR accessory for analysis of biological components and tissues.

Table 2.3. Biological reference samples.

Chemicals	Form	Biological References
Calf thymus DNA	Solid	Nucleic acid
Mucin Type I-S from bovine submaxillary glands	Solid	Protein
Bovine albumin	Solid	Protein
Oleic acid	Solution	Fatty acid
β-carotene Type I - synthetic	Solid	β-carotene (vitamin A)

Table 2.4. Pathological chemicals used for sample preparations.

Chemicals used in Pathology	Form
PBS/D_2O	Solution
OCT	Solution
Formalin	Solution
Paraffin block containing breast biopsy	Solid
Histopure	Solution
Mixture of di-substituted xylene	Solution

2.3.3 FOCAL-PLANE ARRAY FTIR IMAGING SYSTEMS

Infrared images or functional group maps were obtained using a DigiLab Stingray system (DigiLab, Randolph, MA USA), consisting of an FTS7000 series FTIR spectrometer coupled to an UMA600 IR microscope, a Stingray large sample (imaging) accessory and a liquid-nitrogen cooled Lancer™ 64 × 64 MCT focal-plane array (FPA) detector, that collects 4096 infrared spectra simultaneously at the University of Technology Sydney, Sydney. The IR imaging system has a pixel size of 5.5/11 μm with a practical spatial resolution of approximately 15 μm and a field of view (FOV) of 350/700 μm^2 for the 64 × 64 FPA detector [4].

The instrument used DigiLab WinIR Pro version 3.4 Pro, (DigiLab, Randolph, MA, USA) for data collection and manipulation [5]. The instrument was calibrated using a Stingray™ function in the software so that all pixels gave the same apparent response. Backgrounds of 256 co-added scans with a 2 cm^{-1} spectral resolution were collected on a clean surface of a Ag/SnO_2 IR transflective slides. Breast sections were measured using a ×15 magnification and the co-addition of 256 scans with a 2 cm^{-1} spectral resolution.

Spectral data cubes were saved as DigiLab default formats (*.dat extension) [5]. Functional group maps based on the area of diagnostic peaks were plotted. The data cubes were also analysed using Cytospec for multivariate statistical analysis (See Section 2.5). Where appropriate, individual spectra were extracted from the image file and imported into GRAMS/AI™ for further analysis.

2.3.4 SYNCHROTRON-RADIATION BASED FTIR SPECTROMETER

Infrared spectra of breast sections using synchrotron-radiation based FTIR (SR-FTIR) spectroscopy were recorded using a Nicolet Magna 860 FTIR spectrometer (Thermo Nicolet Instruments, Madison, WI, USA) equipped with a Continuμm IR microscope (Spectra Tech, Oak Ridge, TN, USA), mapping stage controller, ×32 objective and MCT detector. The bench was configured with a collimated synchrotron light beamline (BL14A at the National Synchrotron Radiation Research Centre, Taiwan), which served as an external input to the Nicolet Magna 860. The modulated light was directed into the IR microscope for reflection microscopy.

The spectra were collected in the mid-IR range of 4000–700 cm^{-1} at a spectral resolution of 4 cm^{-1} with the co-addition of 32 scans and an aperture of 15 × 15 μm^2. The optics were purged using dry N_2 and the automatic atmospheric suppression function in the OMNIC™ software was used to minimise infrared absorption by CO_2 and water vapour in the ambient air. A single-beam background spectrum was collected from an area free of sample. Stage control, data collection and processing were performed using OMNIC™ version 6.1a (Thermo Nicolet, Madison, WI, USA) [6]. Visible images were obtained using a camera linked to the Continuμm microscope. Experimental data were stored in Thermo Nicolet .map format in OMNIC™ Version 6.1 [6] and false colour functional group maps based on the area of diagnostic peaks were plotted using the Atlμs™ function in OMNIC™ Version 7.1 [7].

2.4 RAMAN SPECTROSCOPY

The samples for Raman spectroscopy and imaging were prepared either as sections or were analysed fresh. Tables 2.5 and 2.6 summarise the experimental protocols for tissue analysis for Raman spectroscopy. For fresh tissue analysis, the specimen with a volume of ~1 mm^3 was placed into a stainless steel sample holder, which was mounted into the macro-sample holder of the FT-Raman spectrometer.

Table 2.5. Experimental protocols for Raman spectroscopic analysis.

Laser Type λ	Ar$^+$: 514 nm	Kr$^+$: 647 nm	Kr$^+$: 752 nm	Diode: 830 nm[#]	Nd:YAG: 1064 nm
Spectrometer	Renishaw Raman Systems 2000 Microscope				Bruker RFS100
Laser power into spectrometer	20–25 mW			≥ 300 mW	130 mW
Laser power at sample	< 2 mW				130 mW (macro)
Objective lenses	×5, ×20 or ×50				n/a
Laser spot size (diameter)	1 μm @ ×50				~100 μm
Grating (lines/mm)	1800	1800	1200	1200	n/a
Spectral region	1800–800 cm^{-1}				3600–50 cm^{-1}
Exposure time	10–30 s (grating spectra)				n/a
Imaging time	600 s (filter image)				n/a
Spectral accumulation	1–5 (grating) and 1 (imaging)				256/512
Laser λ region	Visible		Near infrared		

Table 2.6. Sample preparation for Raman spectroscopy and imaging.

Sample Form	Section Thickness	Laser Excitation Wavelength for Raman Experiments
Bulk tissues in PBS/D_2O	n/a	λ= FT-Raman (1064-nm)
Bulk tissues in formalin	n/a	λ= FT-Raman (1064-nm)
Cryosections on CaF_2/mirror	5–20 μm	λ = 514, 647, 752, 830 and 1064-nm
FFPE sections on CaF_2/mirror	5–20 μm	λ = 514, 647, 752, 830 and 1064-nm
Commercially-available chemical (solid and solutions) standards	n/a	λ = 1064-nm

---

[#] The 830-nm laser was on loan between February and April 2003 from The Medical Laser Institute, The University of Liverpool, UK.

2.4.1 FT-RAMAN SPECTROMETER (λ = 1064-NM)

FT-Raman spectra were collected on a Bruker IFS100 FT-Raman spectrometer (Figure 2.7, Bruker GmbH, Rudolf-Plank-Straβe, Germany) equipped with a neodymium:yttrium aluminium garnet (Nd:YAG) laser emitting at a wavelength of 1064 nm and a liquid-nitrogen-cooled Ge detector with an extended spectral region of 4000–50 cm^{-1}. A 180° back-scattering sampling geometry was employed. Spectra of fresh biopsy samples (see Figure 2.1) and biological reference samples were recorded at a resolution of 4 cm^{-1}, with the co-addition of 256 or 512 scans at a minimum laser power of 130 mW (higher laser powers were also applied to certain samples) and measured in triplicate. The spectra were not corrected for instrument response and all experiments involved the acquisition of spectra in triplicate. A mirror-based quartz cuvette was used in place of aluminium sample holder to measure reference solutions.

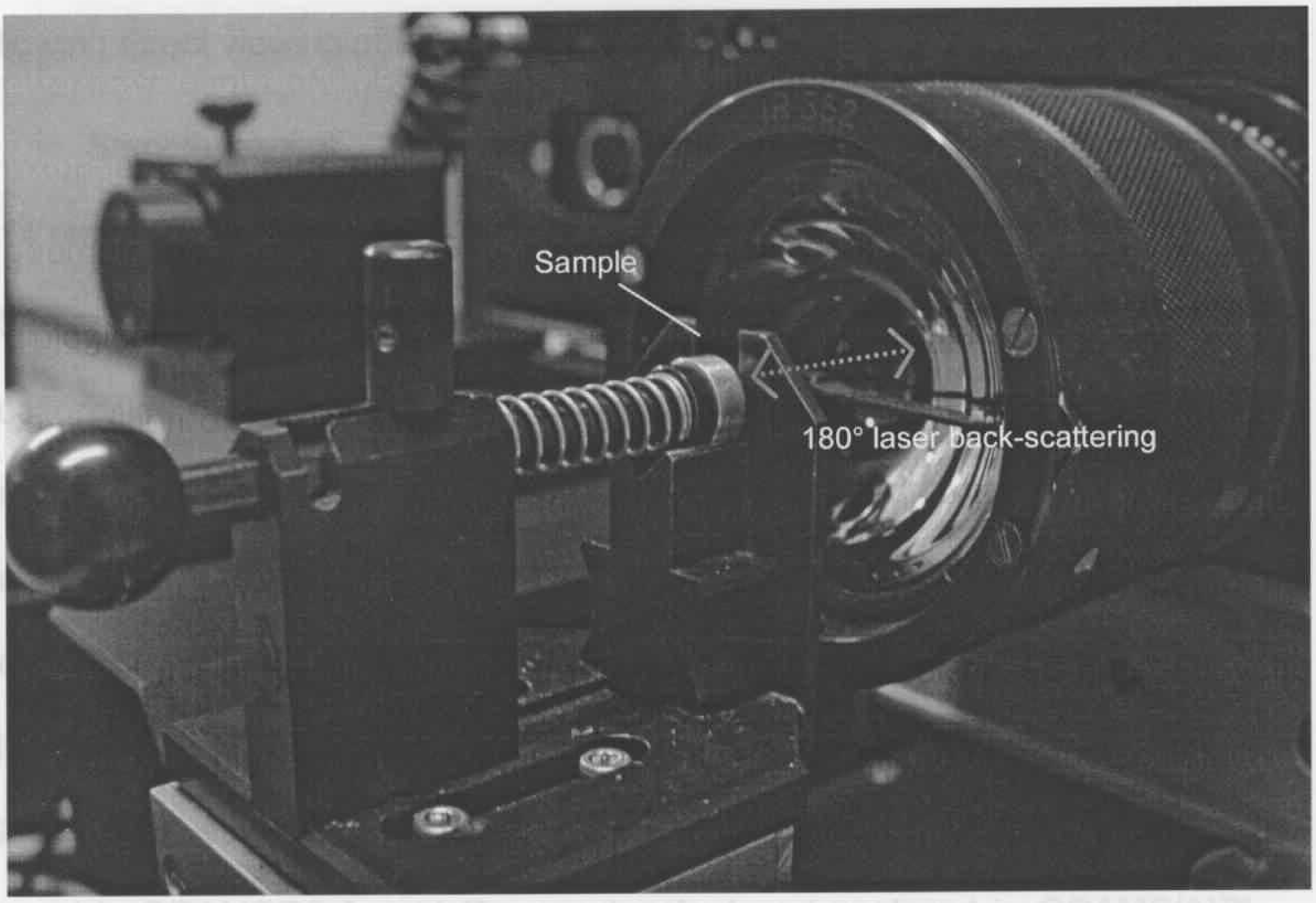


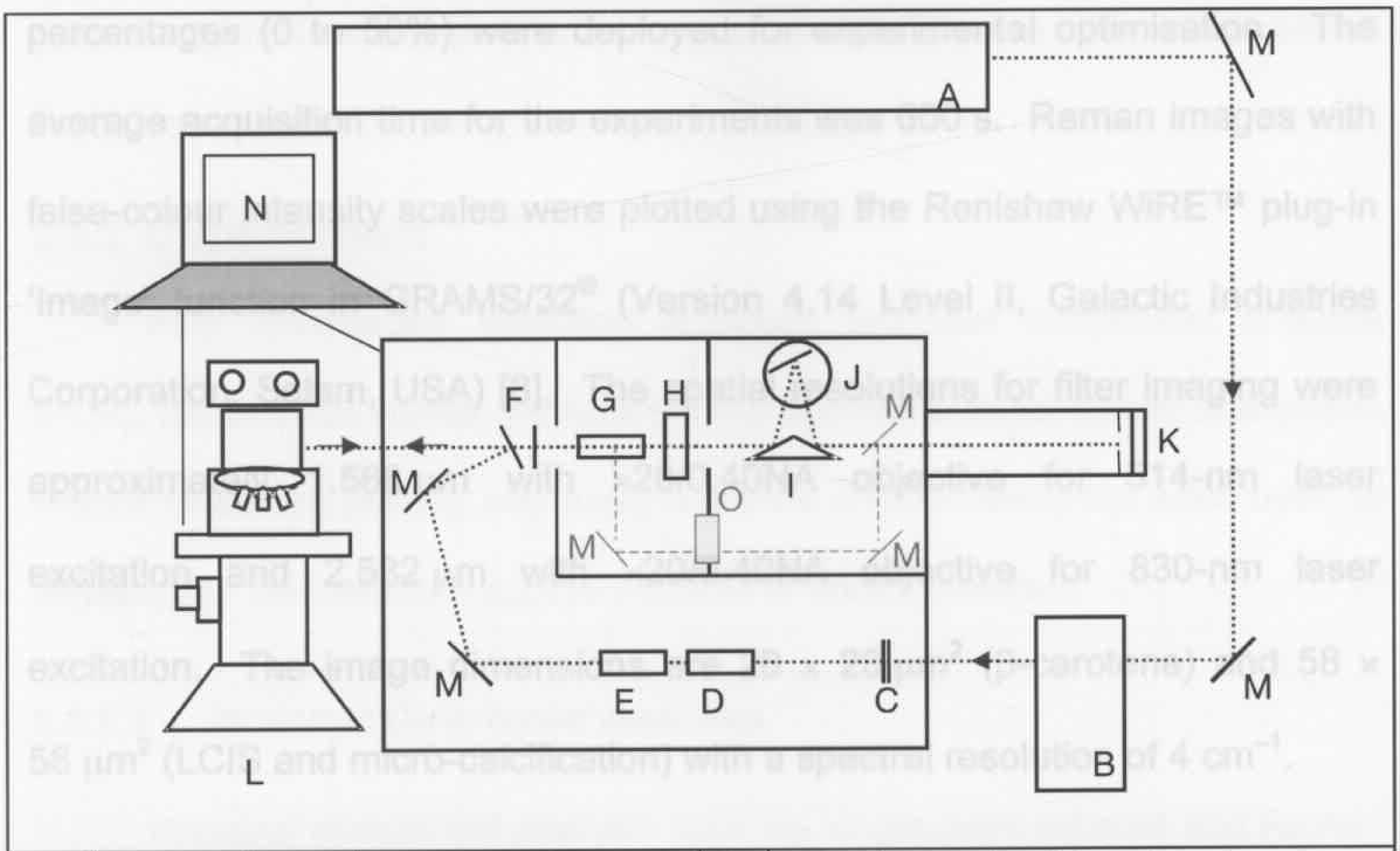
Figure 2.7. FT-Raman spectrometer set-up for bulk tissue analysis, biological references and pathological chemicals analysis.

Experimental data were recorded in triplicate, and saved as Bruker OPUS™ default formats (*.0 extension, version 5.0.53, Bruker Analytik, Germany) and converted to GRAMS™ format (*.spc extension, version 7.02, Thermo Galactic, Salem USA) for spectral analysis in GRAMS™ software [2,3].

2.4.2 DISPERSIVE RAMAN MICROSCOPY: RENISHAW RAMAN SYSTEM

Spectra were collected on a Renishaw Raman Systems 2000 Microscope (Renishaw plc., Wotton-under-Edge, UK), equipped with an air-cooled charge-coupled device (CCD) camera (Figure 2.8). The spectrometer is fitted with holographic notch filters and an 1800 (λ = 514 and 647 nm) or 1200 (λ = 752 and 830 nm) lines/mm grating. The attached microscope is a Leica DMLM equipped with three objectives (×50/0.75 NA, ×20/0.40 NA and ×50/0.12 NA) and a trinocular viewer, which accommodates a video camera allowing direct viewing of the sample on the computer.

Sample excitation was achieved using a Stabilite 2017 Ar ion (λ = 514 nm, Spectra-Physics, CA, USA), a BeamLok 2060 Kr ion (λ = 647 and 752 nm, Spectra-Physics) or a diode laser (λ = 830 nm, Renishaw). Tissue sections were mounted on CaF_2 windows or mirrors and the experimental protocol is summarised in Table 2.5 and spectra were not corrected for instrument response. The spectrometer was computer-controlled using two software packages, one for instrument control (WiRE™ Version 1.3.18, Renishaw) and the other for data analysis (GRAMS/32® Version 4.14 Level II, Galactic Industries Corporation) [8]. Experimental data were acquired and stored in GRAMS/32 format (*.spc extension) and analysed in GRAMS/AI™ Version 7.



A	Kr$^+$ or Ar$^+$ laser source	I	Isosceles triangle mirror
B	Tunable monochromator	J	Grating
C	Computer controlled laser attenuator	K	Charge coupled device (CCD) detector
D	Objective lens	L	Optical microscope
E	Spot focus adjustment lens	M	Adjustable mirrors
F	Plug-and-play holographic notch filter and polariser (wavelength specific)	N	Computer
G	Focus lens	O	Filter wheel for Raman imaging
H	Adjustable slits		

Figure 2.8. Schematic illustration of Renishaw Raman Systems 2000 Microscope with the permission for reproduction from Renishaw plc.

2.4.3 RAMAN IMAGING

Raman images of β-carotene, a lobular carcinoma *in-situ* and a micro-calcification were acquired using a Renishaw Raman Systems 2000 microscope with laser excitations of 514 and 830 nm. The images were collected by deploying filters (the red path shown in Figure 2.8 O) and by using various laser defocus percentages. A filter spectrum of the particular bands of interest was collected prior to the imaging experiment to locate the bands of interests, as the band position may shift by ±10 cm^{-1} when compared to a spectrum collected with the grating. Various laser defocus percentages (0 to 50%) were deployed for experimental optimisation. The average acquisition time for the experiments was 600 s. Raman images with false-colour intensity scales were plotted using the Renishaw WiRE™ plug-in 'Image' function in GRAMS/32® (Version 4.14 Level II, Galactic Industries Corporation, Salem, USA) [8]. The spatial resolutions for filter imaging were approximately 1.568 μm with ×20/0.40NA objective for 514-nm laser excitation and 2.532 μm with ×20/0.40NA objective for 830-nm laser excitation. The image dimensions are 28 × 28 μm^2 (β-carotene) and 58 × 58 μm^2 (LCIS and micro-calcification) with a spectral resolution of 4 cm^{-1}.

2.5 SPECTRAL PROCESSING AND CHEMOMETRICS

2.5.1 SPECTRAL PROCESSING: GRAMS/AI™

Single-point infrared and Raman spectra were converted to the standard .spc format and were analysed primarily in GRAMS/AI™ Version 7. The spectra were processed using various arithmetic programs including smoothing, baseline correction, derivative and curve fitting. Triplicate sample spectra were averaged, and if necessary, baseline corrections and/or smoothing were performed using the second-polynomial Savitzky-Golay method[§] (11 to 17-point). The first and second derivatives were used to determine the width and the position of the peak(s), respectively [9]. The curve-fitting function was applied to determine the positions of the peaks by taking the second derivative and manually fitting the peaks with minimal residual prior to iteration (default = 50 runs). Peak integration was also used to determine the peak area of resolved peaks. Once the spectra were

---

[§] Smoothing function applies to spectra with low signal-to-noise ratio only and the amount used depends on S/N.

processed, multivariate statistical analyses were applied to differentiate breast tissue components.

2.5.2 CHEMOMETRICS

Four multivariate statistical analyses, principal component analysis (PCA), hierarchical cluster analysis (HCA), non-hierarchical cluster analysis k-means clustering (KMC), and fuzzy c-means (FCM) clustering, were used to differentiate the histopathological features of breast lesions.

2.5.2.1 PRINCIPAL COMPONENT ANALYSIS

Principal component analysis from the single-point infrared and Raman spectra were analysed using PCA in The Unscrambler Version 9.6 [10]. Two spectral regions: the C–H stretching region (3050–2750 cm^{-1}) and the 1700–1000 cm^{-1} region, were analysed with full cross-validation and a maximum of twenty principal components. The first two PC scores and the loading were plotted. When outliers were identified from the PC score plots, they were excluded and the PCA was re-analysed.

A PCA image from a DCIS cryosection was analysed using Cytospec Version 1.0.12 [11]. The second derivative of the amide I region (1700–1600 cm^{-1}) was selected for the analysis and the first two PC scores were plotted and imaged.

2.5.2.2 HIERARCHICAL CLUSTER ANALYSIS

Spectra of breast lesions, biological reference samples, and pathology solutions collected using the Bruker FTIR spectrometer were analysed using OPUS hierarchical cluster analysis (HCA). The spectral data were pre-processed using a second-derivative function with nine-point smoothing and/or vector normalisation. The distance matrix was calculated using D-value over the selected spectral regions, followed by applying the Ward's cluster algorithm. The dendrogram was produced to differentiate various components [2].

Two HCAs were applied: (1) identification of unknown breast lesions from the representative spectra of known benign and malignant breast lesions; (2) differentiation of specific breast lesion based on the IR bands of biological components (proteins, lipids and nucleic acids).

2.5.2.3 NON-HIERARCHICAL CLUSTER ANALYSIS

Non-hierarchical cluster images using k-means clustering (KMC) and fuzzy c-means (FCM) clustering algorithms on the IR images of breast sections were applied to differentiate various breast tissue components in Cytospec [11]. In KMC analysis, the spectral regions of 3800–3000, 3000–2800, 1800–950, 1750–1580, and 1300–1000 cm^{-1}, together with two, three, five and/or six clusters, and 100 (default) training cycles were applied. The same spectral regions with minimum improvement value of 0.0005 (default) were applied in the FCM cluster analysis using the normalised data with five clusters to generate the individual cluster plots. In both non-HCA, the average spectra of each cluster were also acquired.

2.6 REFERENCES

[1] J. Radford, *Histopathology Techniques.* 7th ed., The University of Sydney, Sydney, 2002.

[2] Bruker Optik, *Software: Bruker Opus Version 5.0.53,* 2003.

[3] Thermo-Informatics, *Software: Grams/AI Version 7.02,* 2002.

[4] Varian, Lancer™ Focal Plane Array (FPA) Cameras. http://www.varianinc.com/cgi-bin/nav?products/spectr/ftir/ftir_imaging/lancer_fpa&cid=INQMILPMFI, Varian, 2005.

[5] Varian, *Software: Digilab Win-IR Pro Version 3.4,* 2003.

[6] Thermo Nicolet, *Software: Omnic™ 6.1a,* 2001.

[7] Thermo-Nicolet, *Software: Omnic™ Atlμs™ 7.1,* 2005.

[8] Thermo-Informatics, *Software: Grams/32 Version 4.14 Level II,* 1999.

[9] B. Stuart, W. George, P. McIntyre, *Modern Infrared Spectroscopy: Analytical Chemistry by Open Learning*. John Wiley & Sons, Chichester, 1996.

[10] CAMO Process AS, *Software: The Unscrambler, Version 9.6*; 2006.

[11] Cytospec, *Software: Cytospec, Version 1.2*; 2005.

3. HISTOPATHOLOGY OF BREAST TISSUES

3.1 BREAST PATHOLOGY REPORTS

A total of 100 breast biopsies were collected with the ethics clearance approvals. These samples were diagnosed by pathologist* as follows: 43 benign (36 fibrocystic changes, 4 fibroadenoma and 3 intraduct papilloma); 7 *in-situ* carcinomas (5 DCIS and 2 LCIS); and 47 malignant breast cancers (36 IDC and 11 ILC), see Section 2.1 (Table 3.1). Three samples were unable to be classified due to insufficient sample size and a lack of morphological features. No normal breast samples were obtained for this research as the samples were acquired either from open-core or surgical biopsies upon the initial results from triple assessment with approval from Human Ethics Committees. Ten samples acquired were not sent to pathology prior to spectroscopic analysis; these were used for a blind study. The majority of the biopsies for histopathology were sectioned using formalin-fixed paraffin-embedded (FFPE) technique and to a lesser extent, cryostat sectioning. Sections prepared from both techniques were stained with haematoxylin and eosin (H&E), as described in Section 2.2, and were classified using light microscopy by a pathologist [1].

Table 3.1. Pathological results on breast specimens (100).

Groups	Classification*
Benign	**43**
Fibrocystic changes	*36*
Fibroadenoma	*4*
Intraduct Papilloma	*3*
***In-situ* Carcinomas**	**7**
DCIS	*5*
LCIS	*2*
Malignant	**47**
IDC	*36*
ILC	*11*
Unclassified	**3**

* Pathologist: Professor Peter Russell and early samples were pre-identified upon sample collection from The Institute for Magnetic Resonance Research, The University of Sydney.

DCIS: ductal carcinoma *in-situ*; LCIS: lobular carcinoma *in-situ*; IDC: invasive ductal carcinoma; and ILC: invasive lobular carcinoma.

3.2 NORMAL BREAST FEATURES

Since no normal breast specimens were collected for this research, lesion-free areas in the breast specimen were used as 'normal' controls. These areas were designated as lesion-free by the pathologist. Figures 3.1 and 3.2 are representative examples of the major breast tissue components of the H&E stained 'normal' breast sections.

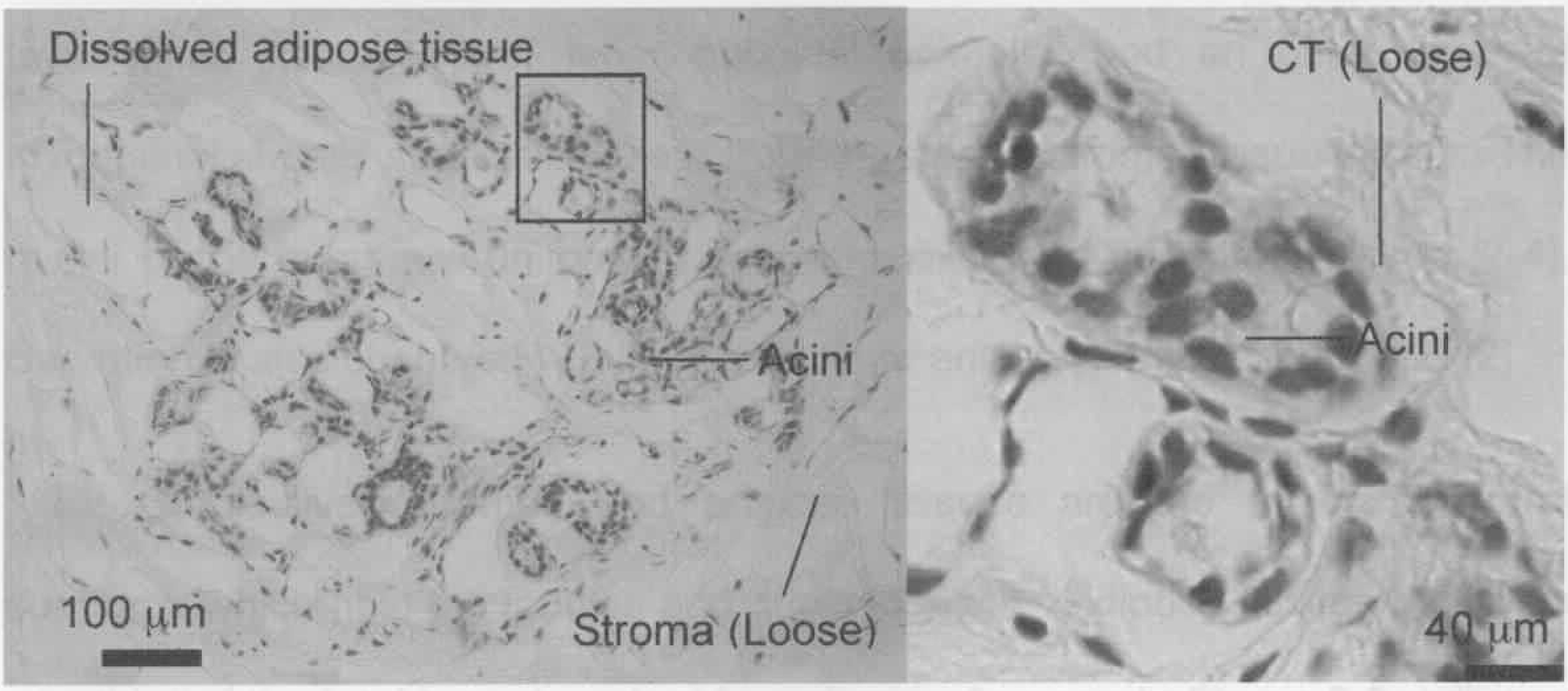


Figure 3.1. H&E of a FFPE section of normal lobules. The nuclei of the acini are uniform in shape and they are surrounded by loose stromal tissues and adipose tissues (dissolved). (Left) low magnification and (right) high magnification showing the acini. CT: connective tissue.

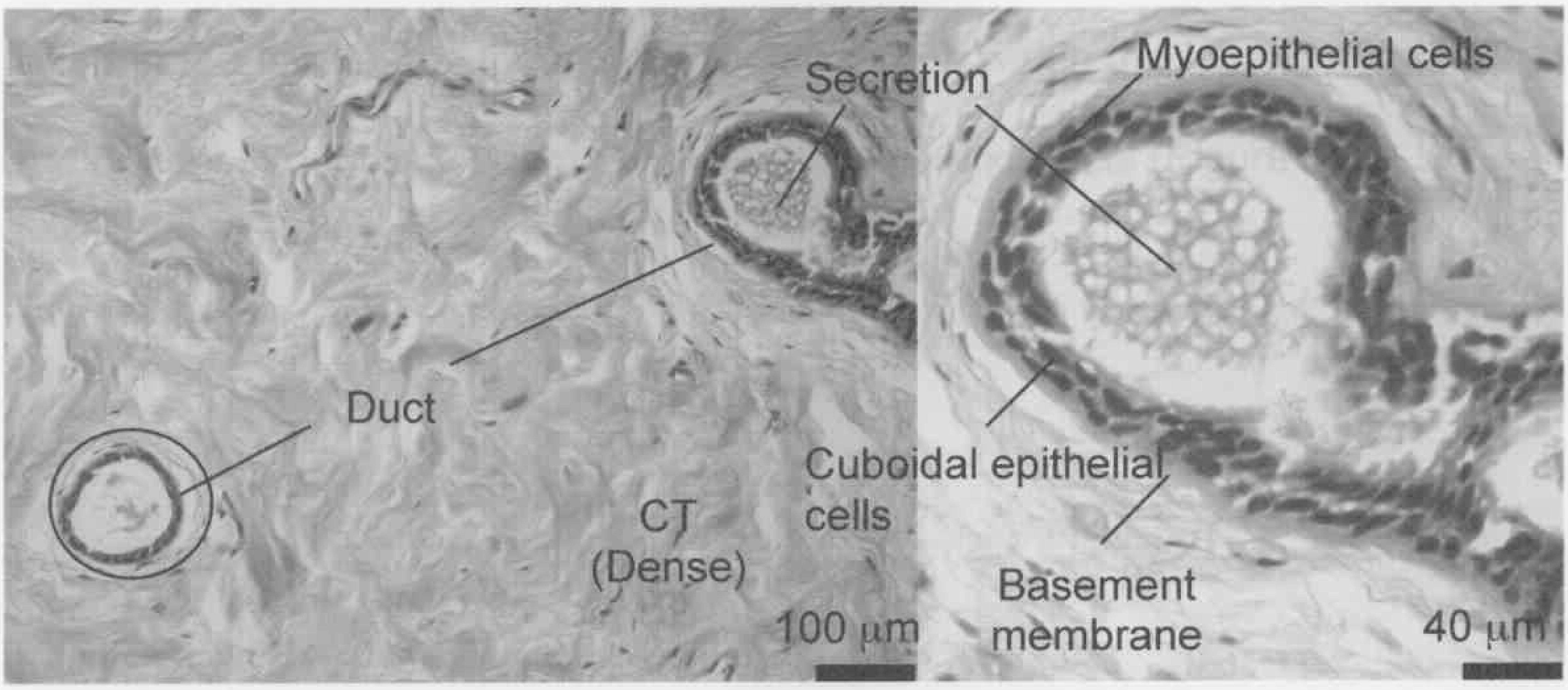


Figure 3.2. H&E of a FFPE section of normal ducts. The ducts here are surrounded by dense stromal tissues. The ducts are classified as normal showing an inner layer of cuboidal epithelial and outer myoepithelial cells. (Left) low magnification and (right) high magnification showing the epithelial components of the duct with secretion.

The lobules (Figure 3.1) are comprised of a collection of milk production units called acini; each acinus is composed of an inner layer of cuboidal epithelial and an outer layer of myoepithelial cells that are surrounded by loose stromal tissue and adipose tissue (dissolved). Closer examination of the acini illustrates the interlobular connective tissues within the lobules [2].

The ducts in Figure 3.2 were classified as normal as they show the two-layer lining composed of inner cuboidal epithelial and an outer layer of myoepithelial cells separating them from the dense stromal tissues [2-4]. The nuclei of the ducts are uniform in shape as expected for normal tissue [2-4]. Secretion is also observed within the duct (see enlargement of Figure 3.2) [2].

Connective (stromal) and adipose tissues are the major structural supports, which hold the lobules and ducts in place. Adipose tissues comprise one third of the total breast content by volume. As seen in Figure 3.3, adipose tissues appear as openings as a result of the adipose being dissolved in ethanol during tissue processing. Two types of stromal tissues are commonly found in the breast: the loose intralobular stroma within the lobules (Figure 3.1) and denser interlobular stroma between lobules and ducts (Figure 3.3).

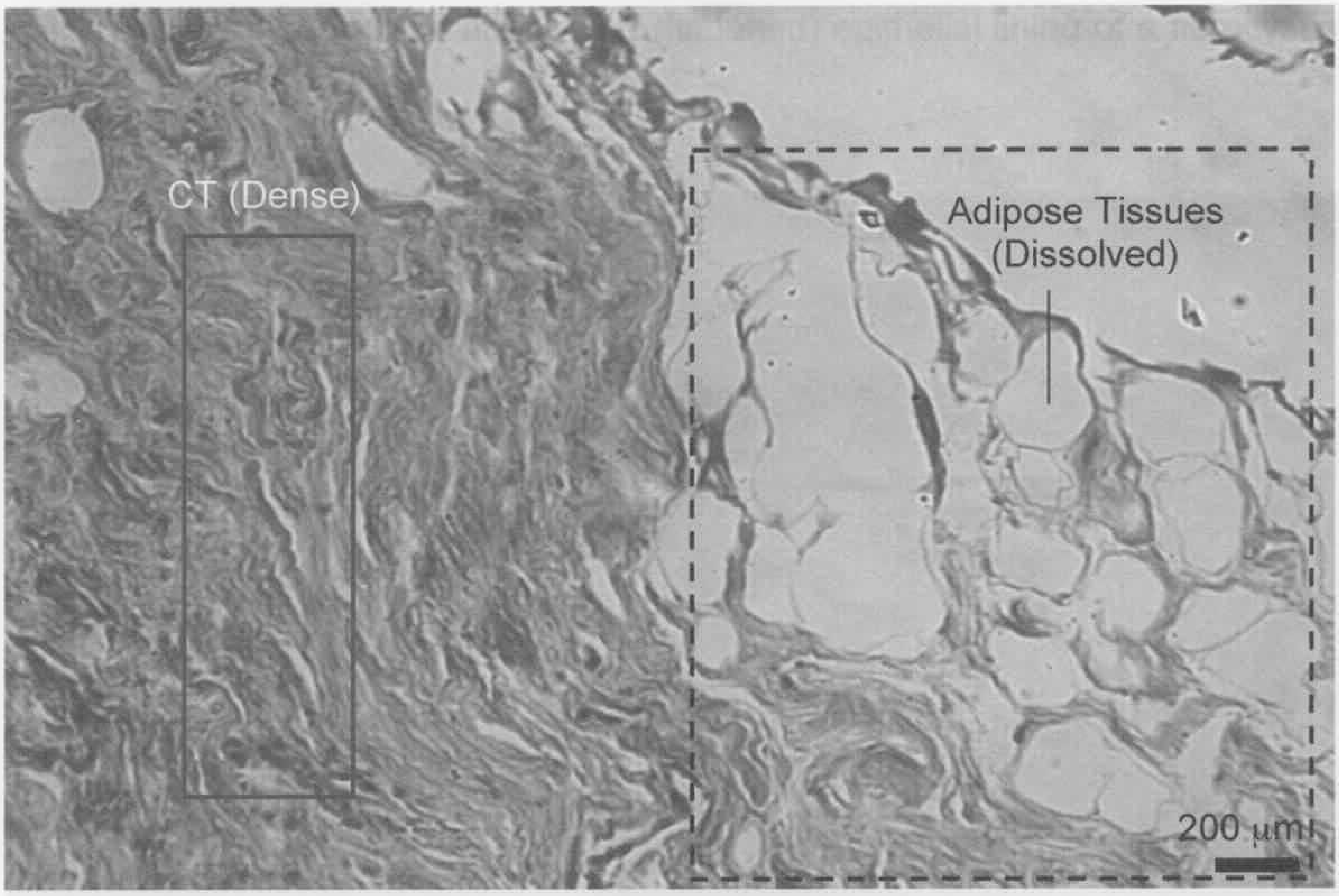


Figure 3.3. H&E of a FFPE section of connective and adipose tissues at low magnification. The dense connective tissues seen here support the individual breast tissue components. Adipose tissues are dissolved during tissue processing and appear as openings, which do not contain any epithelial cells as seen in the ducts and lobules.

3.3 BENIGN BREAST DISEASES

Benign breast tumours are not cancerous since they do not spread to other locations in the body [4]. The benign conditions diagnosed in this study are fibrocystic changes (36 cases), fibroadenoma (4 cases) and intraduct papilloma (3 cases) [2,3].

3.3.1 FIBROCYSTIC CHANGES

Fibrocystic changes are commonly found in women, and account for 83% of all benign lesions studied in the following sections. The histological features of fibrocystic changes are enlargement of the breast duct due to the formation of fluid-filled sacs called cysts (Figures 3.4–3.6) and the absence of tumour cells inside the duct or basement membranes [2-4]. The sizes of the cysts are variable, resulting in different sized fibrocystic ducts as seen in Figure

3.4. Figure 3.5 shows the attenuated (flattened) epithelial lining of a fibrocystic duct [2].

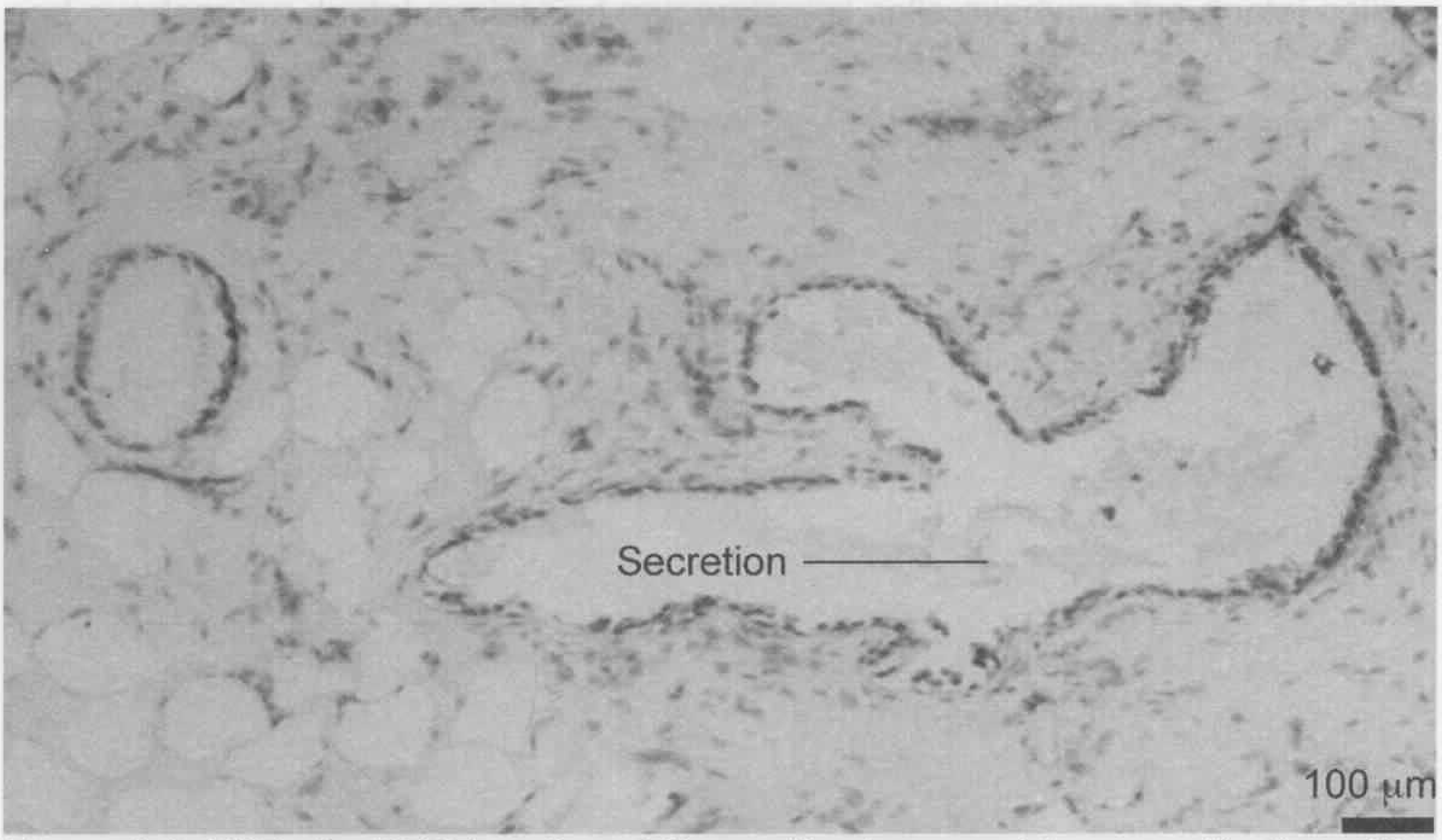


Figure 3.4. H&E of a FFPE section of fibrocystic changes at low magnification. The duct on the right appears enlarged and secretion is found within the duct, which is the common characteristic of fibrocystic changes.

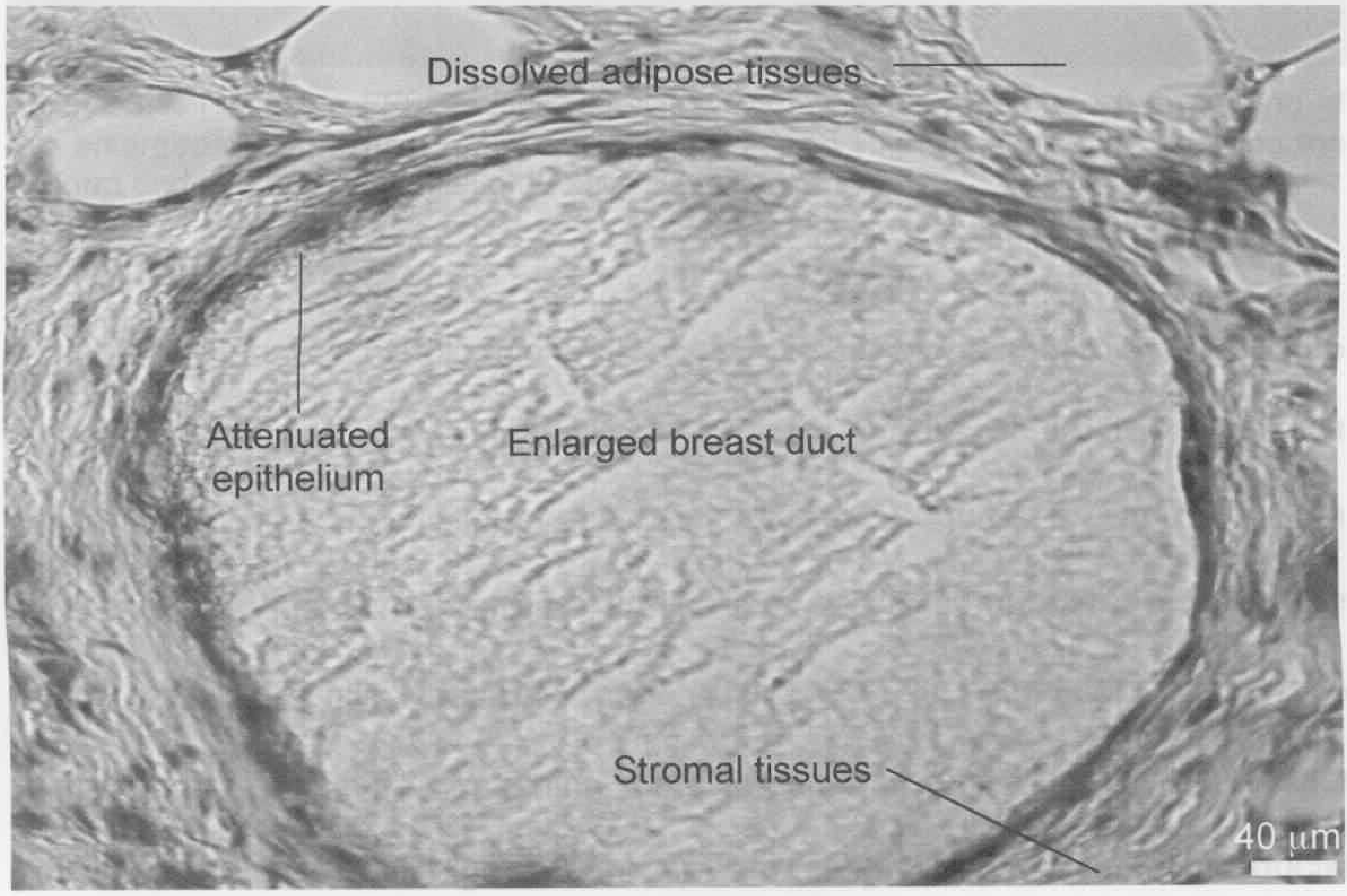


Figure 3.5. H&E of a FFPE section of a fibrocystic change at high magnification showing the enlarged breast duct and the attenuated ductal epithelial cells containing secretion and surrounded by adipose tissue and stroma.

Figure 3.6 depicts two merging fibrocystic ducts. As opposed to the attenuated epithelial structure in Figure 3.5, these two ducts contain uniform and overlapping ductal nuclei and myoepithelial cells overlying the ductal cells.

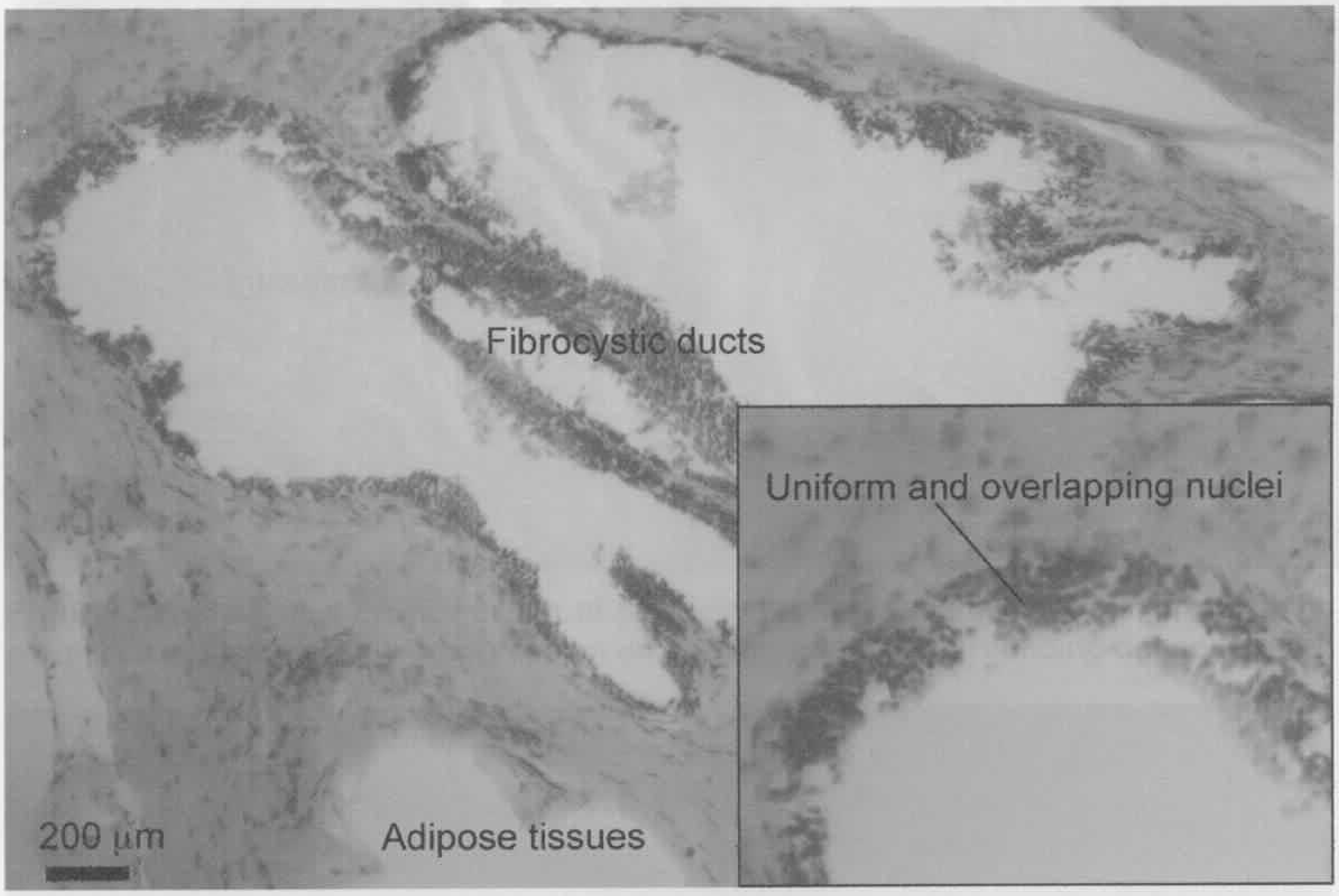


Figure 3.6. H&E of a FFPE section showing fibrocystic changes and the merging of two enlarged ducts. The stroma are normal. (Inset) High magnification showing the uniform and overlapping nuclei of the fibrocystic duct.

3.3.2 FIBROADENOMA

There are two histological patterns commonly found in fibroadenoma: the intracanalicular and the pericanalicular patterns. Four of the fibroadenoma samples exhibited mostly intracanalicular patterns (Figures 3.7 and 3.8). In such patterns, the fibrous stroma proliferates and is compressed around the ductal spaces while the basement membrane is left intact [2,5]. A fibroadenoma section, containing both intracanalicular and pericanalicular patterns, is depicted in Figure 3.8. The stroma proliferate around the ductal spaces in pericanalicular fibroadenoma resulting in a round or oval cross-section [2-4].

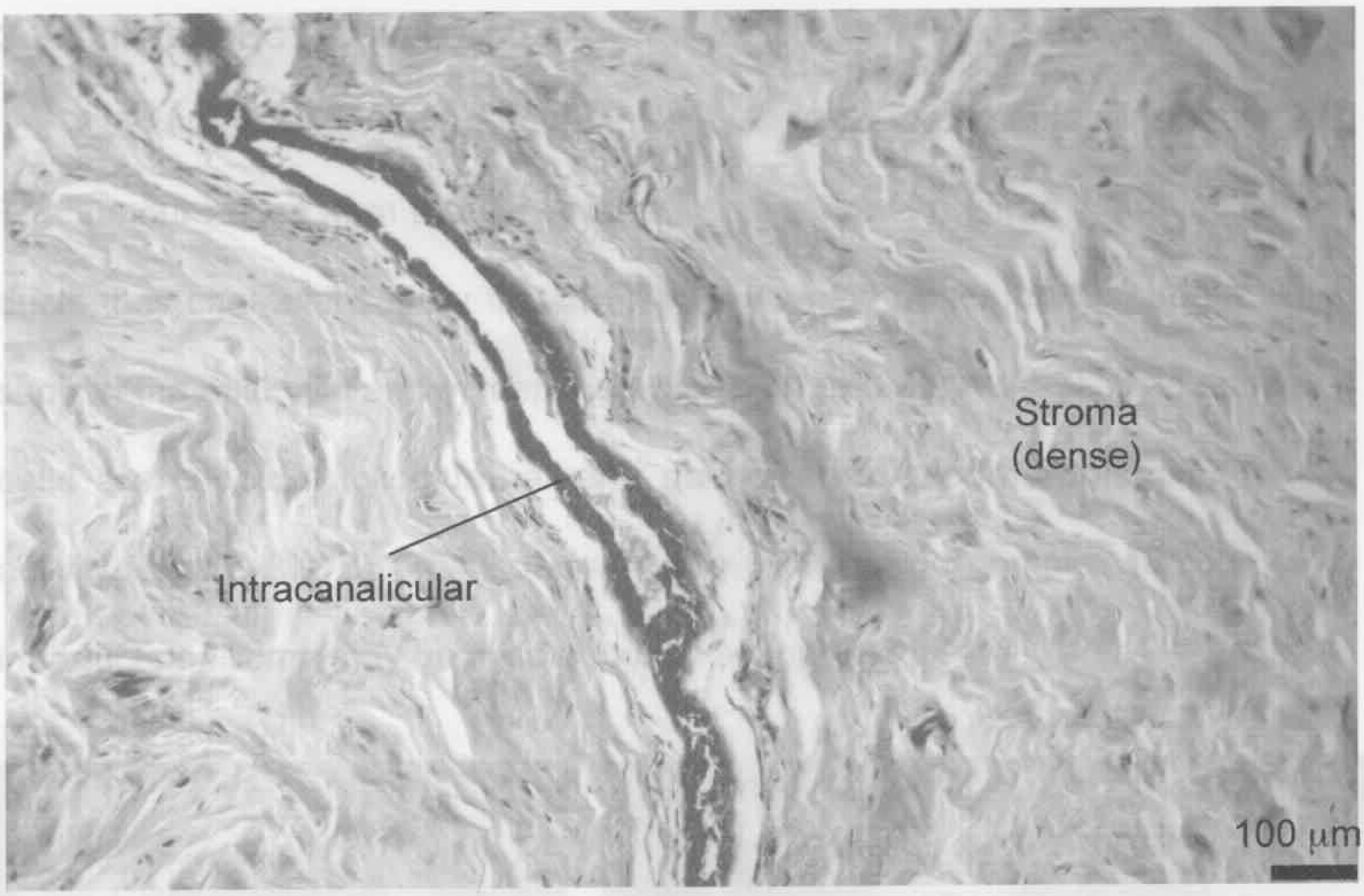


Figure 3.7. H&E of a FFPE section of fibroadenoma (intracanalicular pattern) at high magnification showing the compression of the duct from the surrounding dense stroma.

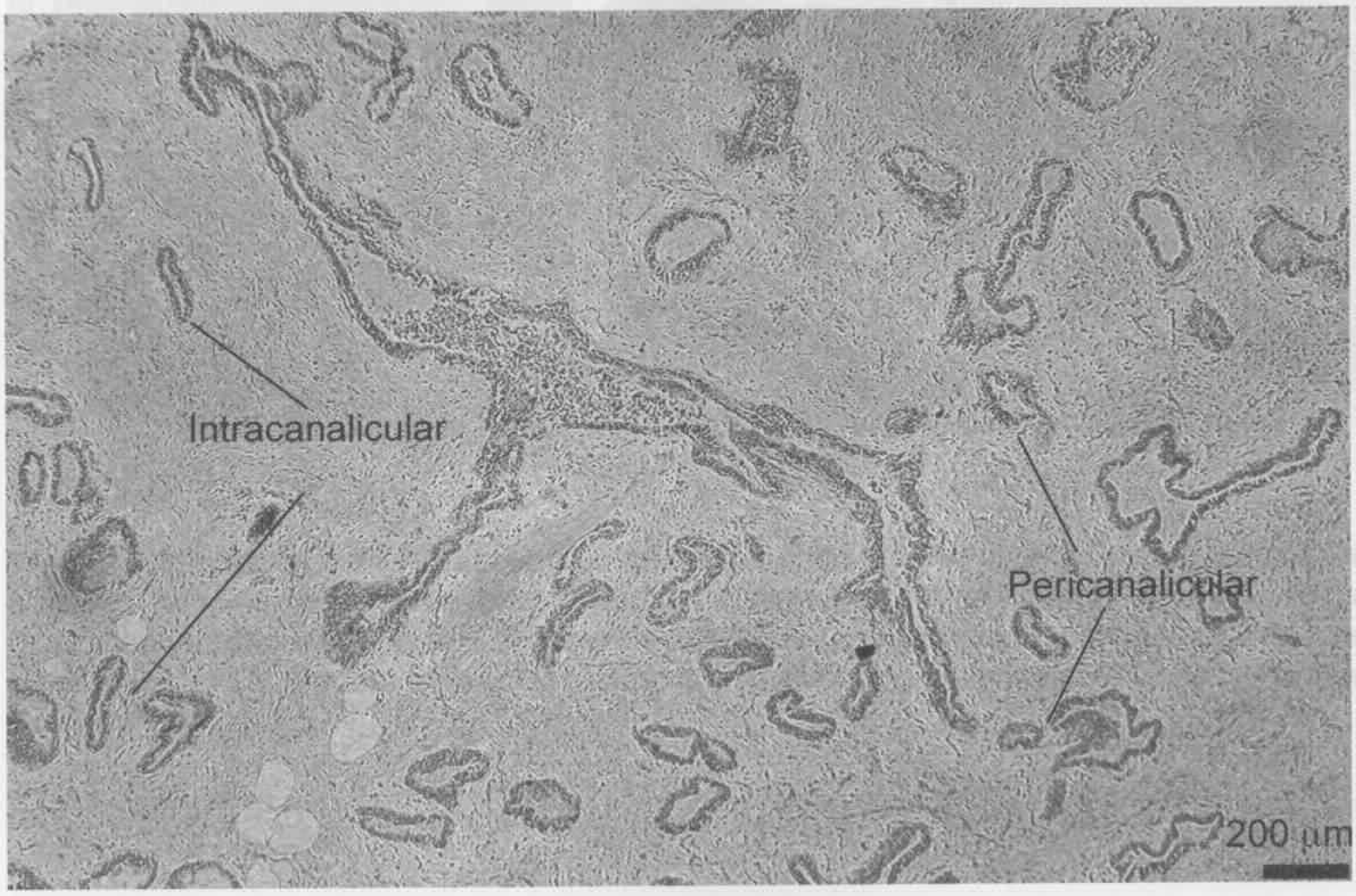


Figure 3.8. H&E of a FFPE section of a fibroadenoma (low magnification) showing both pericanalicular and intracanalicular patterns. The compressed elongated ducts are intracanalicular patterns with the duct being compressed. The round and oval duct cross-section are that of pericanalicular pattern. The large duct contains suspicious tumour cells and is also compressed.

3.3.3 INTRADUCT PAPILLOMA

A representative intraduct papilloma section is presented in Figure 3.9. The tumour arises from the duct epithelium and is often found in the lactiferous ducts, it is characterised by the complex branching fronds of the stroma which support the epithelial and myoepithelial cells (Figure 3.9 A) [4-7]. The epithelial cells are cuboidal to columnar and do not show pleomorphism or mitotic activity (Figure 3.9 B) [7]. However, due to the large complex tumour cells, intraduct papilloma may be misdiagnosed as malignant [4-7].

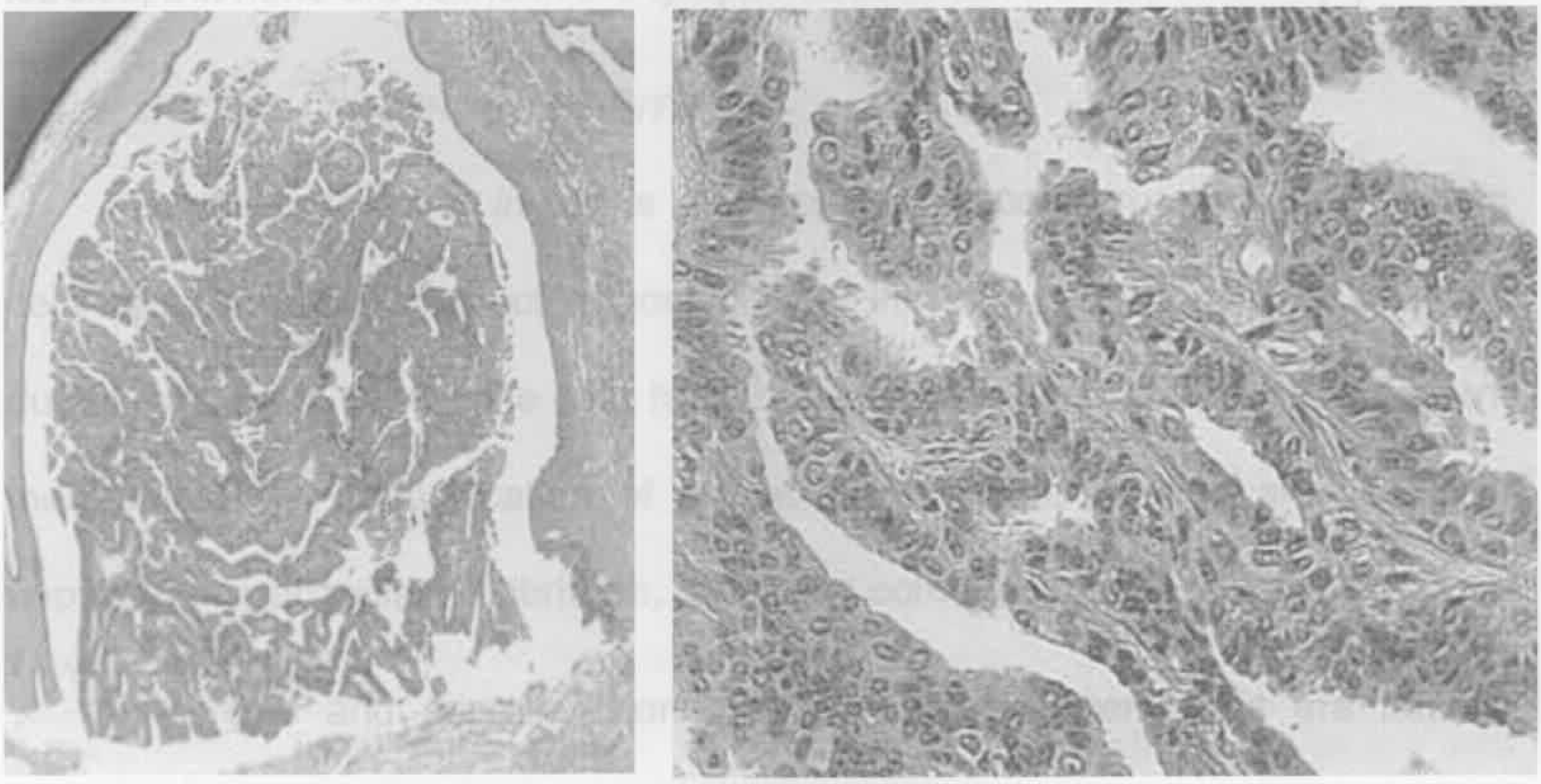

(A) (B)

Figure 3.9. H&E of a FFPE section of an intraduct papilloma at: (A) low magnification showing the complex branching fronds within a lactiferous duct; and (B) high magnification showing the two-cell layer epithelium. Images were reproduced with the permission of Dr. Donald Innes, Department of Pathology, University of Virginia Health Sciences Center [7].

3.4 *IN-SITU* CARCINOMAS (TNM: T_{IS}, N_0, M_0)

In-situ means 'in place'; there are two types of *in-situ* carcinomas: ductal carcinoma *in-situ* (DCIS) and lobular carcinoma *in-situ* (LCIS) [4]. The tumour cells from both DCIS and LCIS do not spread outside the ducts or lobules. They can develop into invasive breast cancer or increase the risk of developing invasive breast cancer, if not treated properly [8]. According to the TNM staging system for breast cancer classification (Section 1.3), both DCIS and LCIS are categorised as T_{is}, N_0 and M_0, as tumour cells are found within the ducts or lobules, but none are found in the lymph nodes or have metastasised [9].

3.4.1 DUCTAL CARCINOMA *IN-SITU*

Ductal carcinoma *in-situ* is a neoplastic process in which proliferating cells remain within the duct without cellular invasion, as opposed to invasive ductal carcinoma, where the cells have breached the boundary of the basement membrane [2]. Classification of the lesion is based on five architectures: papillary, micropapillary, cribriform, solid, and comedo [10-12].

Necrosis and microinvasion of the basement membrane are usually found in the comedo subtype (high-grade DCIS), which has a higher risk of developing invasive breast cancer [10,12]. High-grade DCIS also displays linear branching or coarse granular calcification, formation of abnormal nucleoli and chromatin and clusters of large tumour cells. Low-grade DCIS, such as cribriform DCIS, often shows fine granular calcification similar to benign lobular calcification, the absence of tubules, stromal infiltration by tumour cells and clusters of mildly pleomorphic malignant cells [4,6].

An example of low-grade DCIS in the cribriform architecture is depicted in Figure 3.10. The sieve-like features are clearly seen here as the tumour cells are located within the breast duct without invasion to the basement membranes. The tumour cells in the cribriform DCIS are small and uniform compared to normal epithelial cells (Figure 3.1). The cribriform ducts are surrounded by loose stromal tissues. The intracellular connective tissues hold the tumour cells within the duct and there is no apparent cellular invasion to the basement membranes.

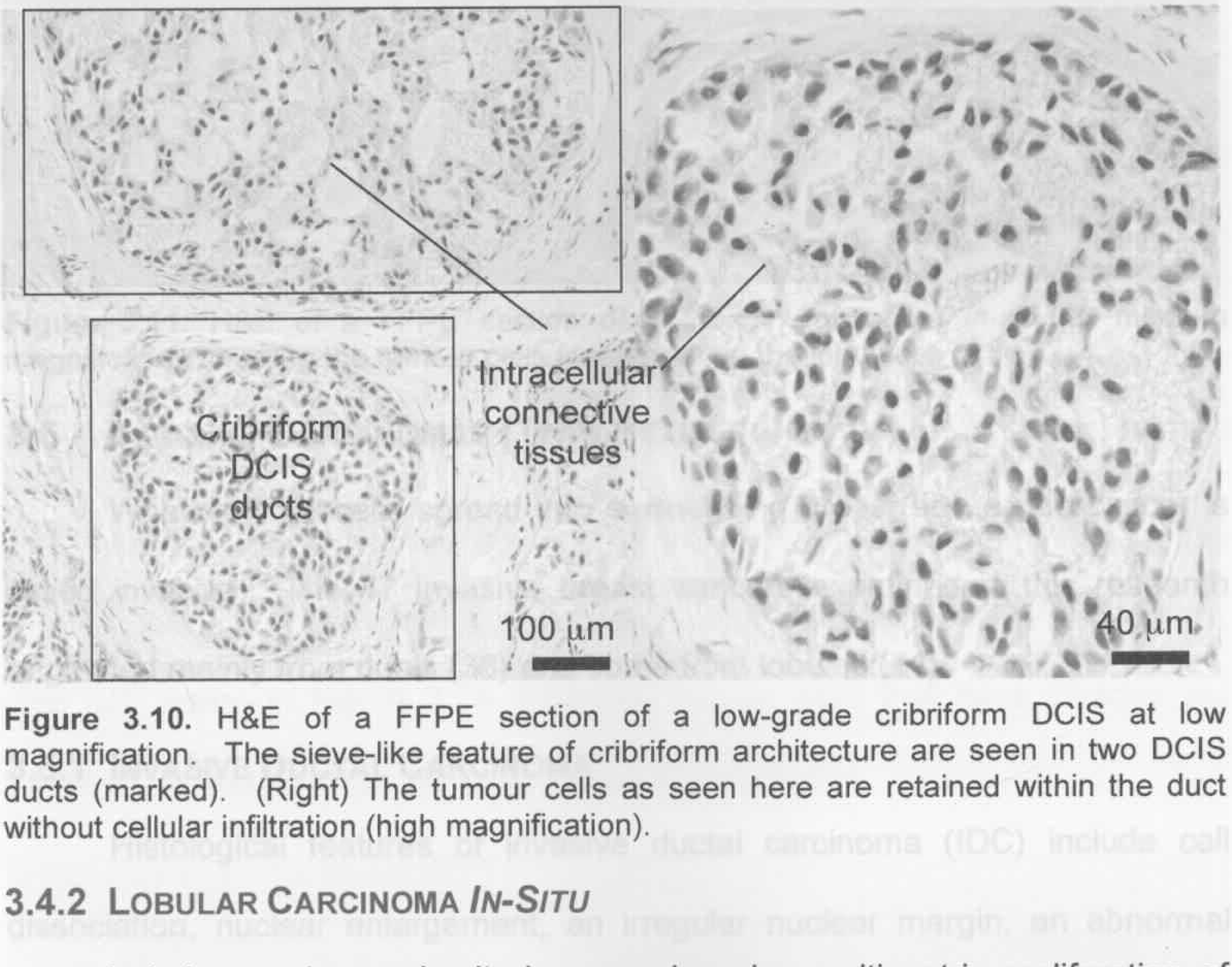


Figure 3.10. H&E of a FFPE section of a low-grade cribriform DCIS at low magnification. The sieve-like feature of cribriform architecture are seen in two DCIS ducts (marked). (Right) The tumour cells as seen here are retained within the duct without cellular infiltration (high magnification).

3.4.2 LOBULAR CARCINOMA *IN-SITU*

Lobular carcinoma *in-situ* is a non-invasive multicentric proliferation of atypical epithelial cells in the lobules and terminal ducts of the breast [13]. It is not usually detected by mammography [4] but two LCIS cases within the 100 biopsies examined were diagnosed by pathology. It is usually found with other invasive carcinomas, which makes diagnosis difficult [4,6,13]. In Figure 3.11, the lobules are filled with a uniform population of small tumour cells (arrows).

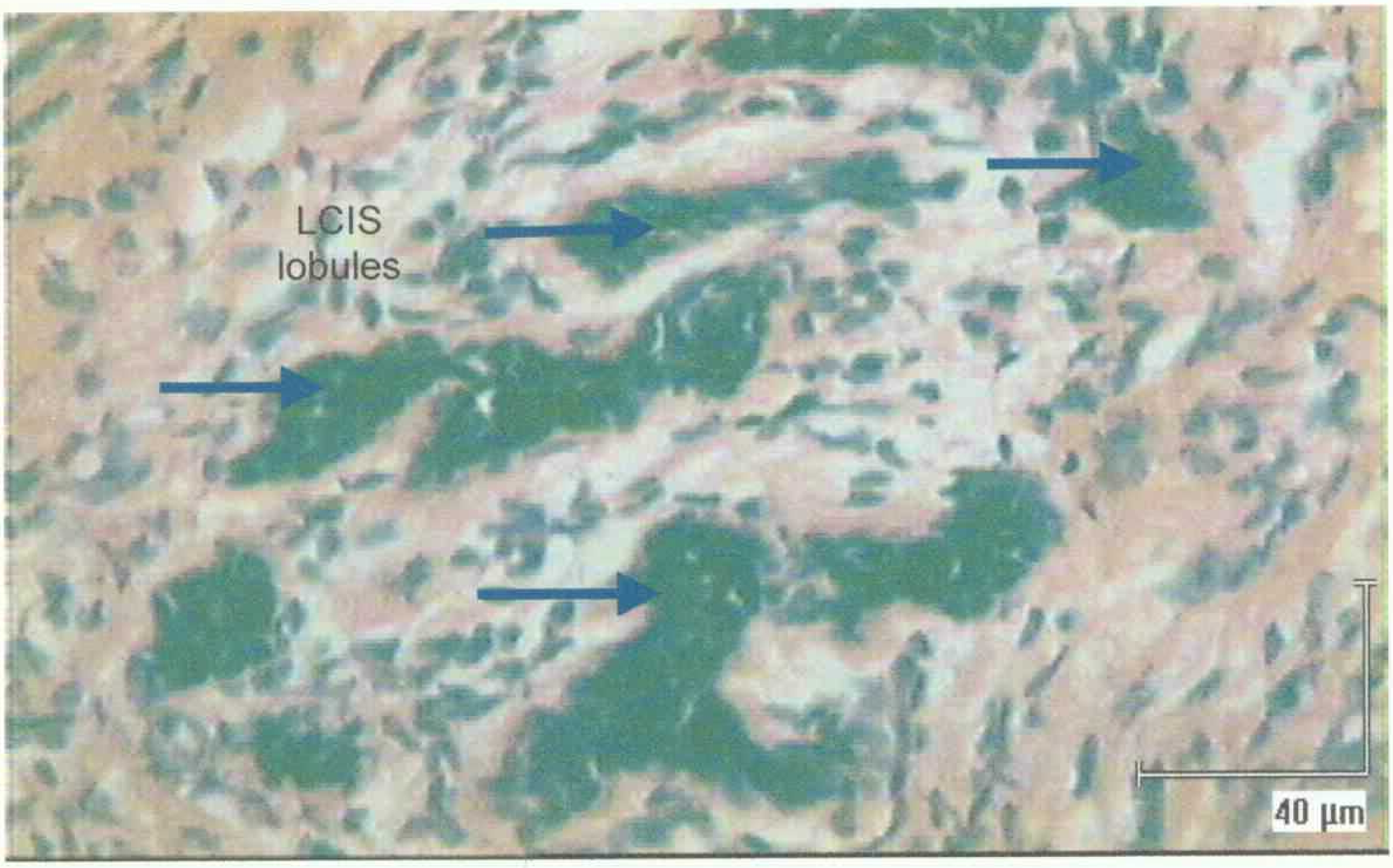


Figure 3.11. H&E of a FFPE section of a lobular carcinoma *in-situ* at medium magnification showing the tumour cells located within the lobules.

3.5 INVASIVE CARCINOMAS (TNM STAGES: I TO IV)

When tumour cells spread into surrounding breast tissue, the cancer is called invasive. The 47 invasive breast cancers examined in this research originated mainly from ducts (36) and some from lobules (11).

3.5.1 INVASIVE DUCTAL CARCINOMA

Histological features of invasive ductal carcinoma (IDC) include cell dissociation, nuclear enlargement, an irregular nuclear margin, an abnormal chromatin pattern, nucleoli, mitoses and intranuclear inclusions [2,4,6]. Figure 3.12 shows an IDC lesion, in which there are tumour cells within the ducts and cellular invasion is apparent through basement membranes to the surrounding breast tissue. Calcification is also found in one of the ducts. Most of the calcifications from malignant lesions are in the form of calcium hydroxyapatite, whereas calcium oxalate dihydrate is usually found in benign lesions [2,4].

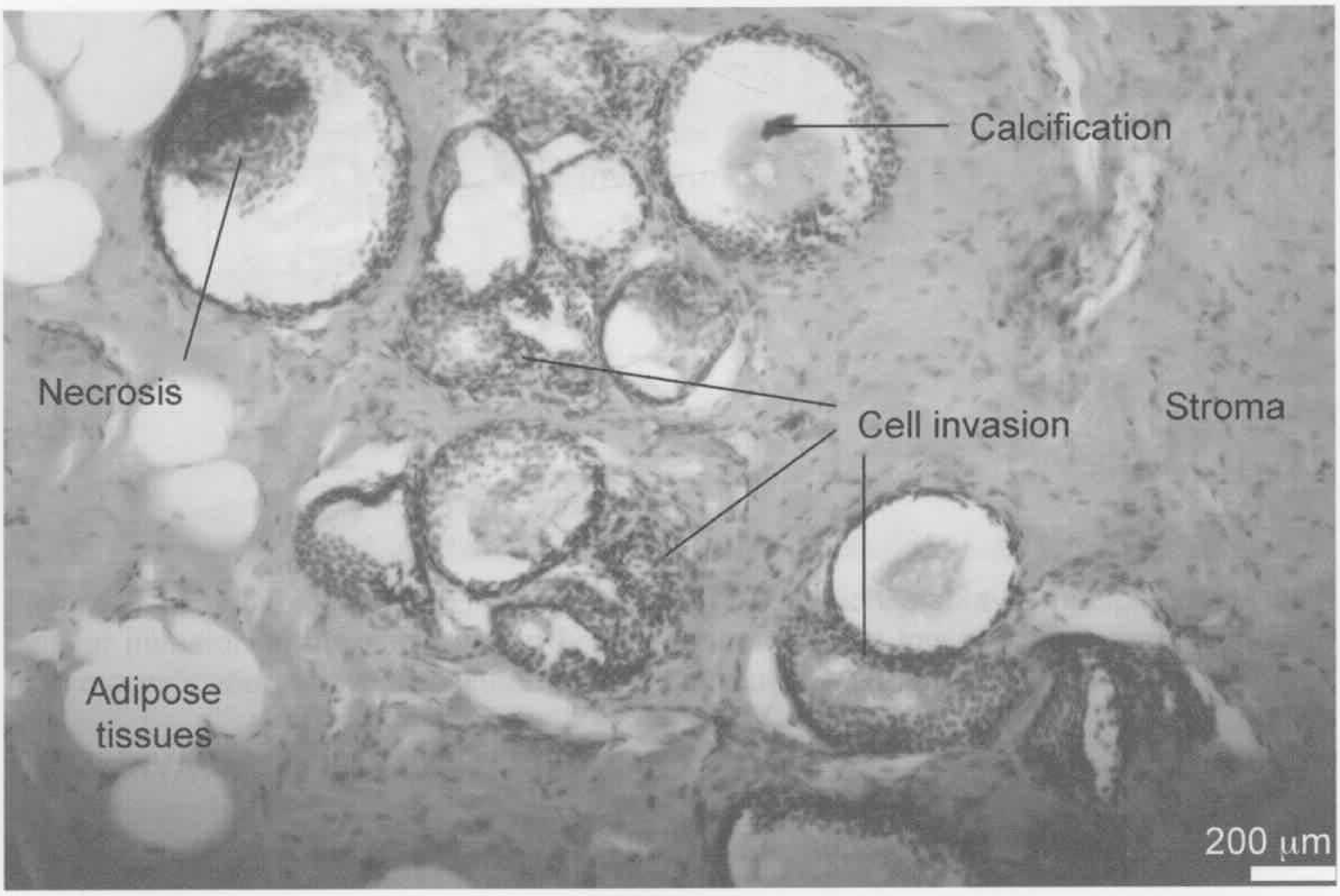


Figure 3.12. H&E of a FFPE section of an invasive ductal carcinoma at low magnification showing the cellular infiltration to the surrounding breast tissues. The ducts shown here are in the comedo form, which contains multiple epithelial cells and necrosis.

Cellular invasion from the duct is clearly identified in Figure 3.13; the tumour cells are uniform and lymphocytes due to cell inflammation are present. The IDC lesion in Figure 3.14 illustrates tumour cells that have invaded the basement membranes to the surrounding breast tissues.

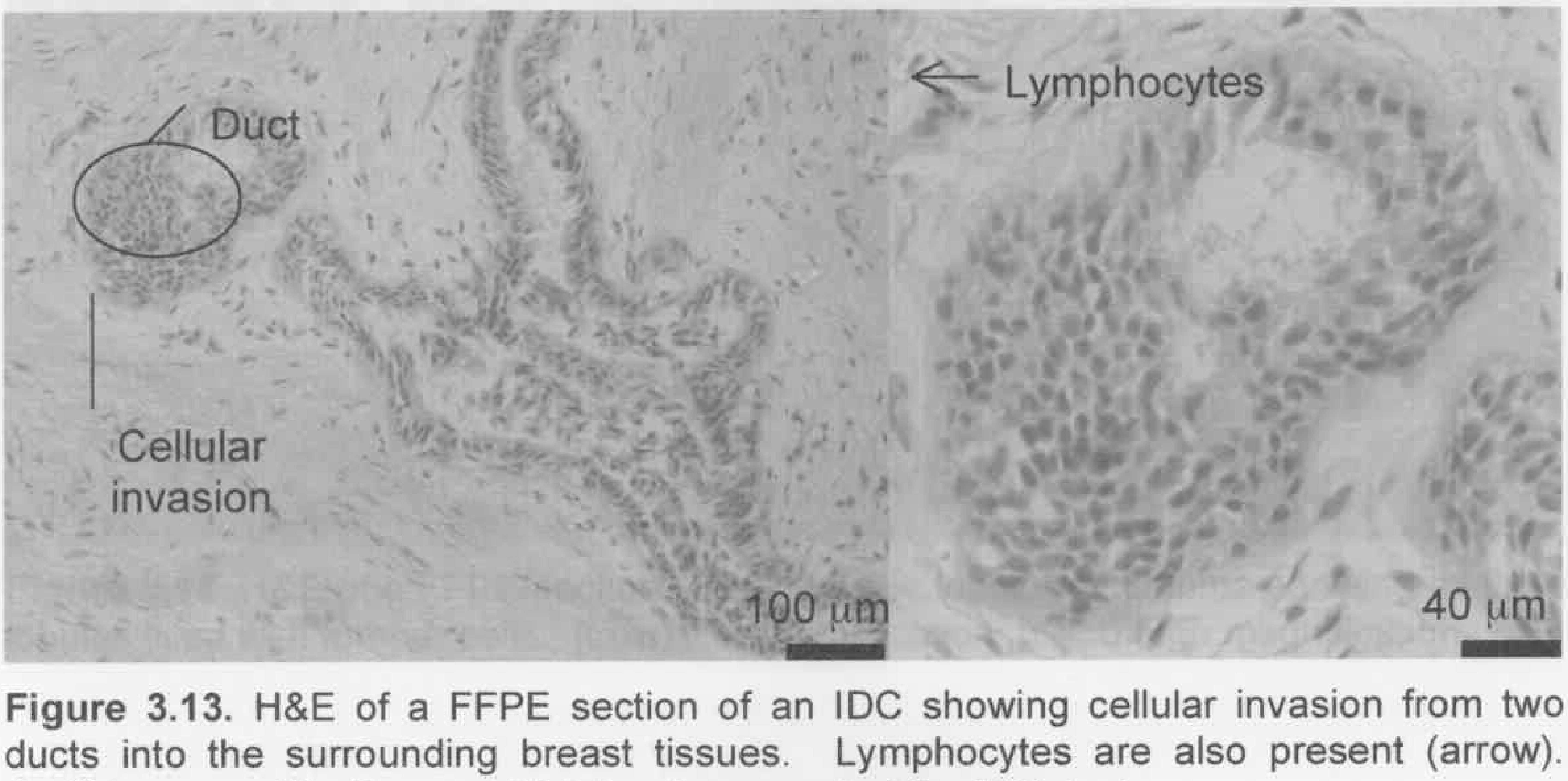


Figure 3.13. H&E of a FFPE section of an IDC showing cellular invasion from two ducts into the surrounding breast tissues. Lymphocytes are also present (arrow). (Left) low magnification and (right) enlargement of the IDC duct.

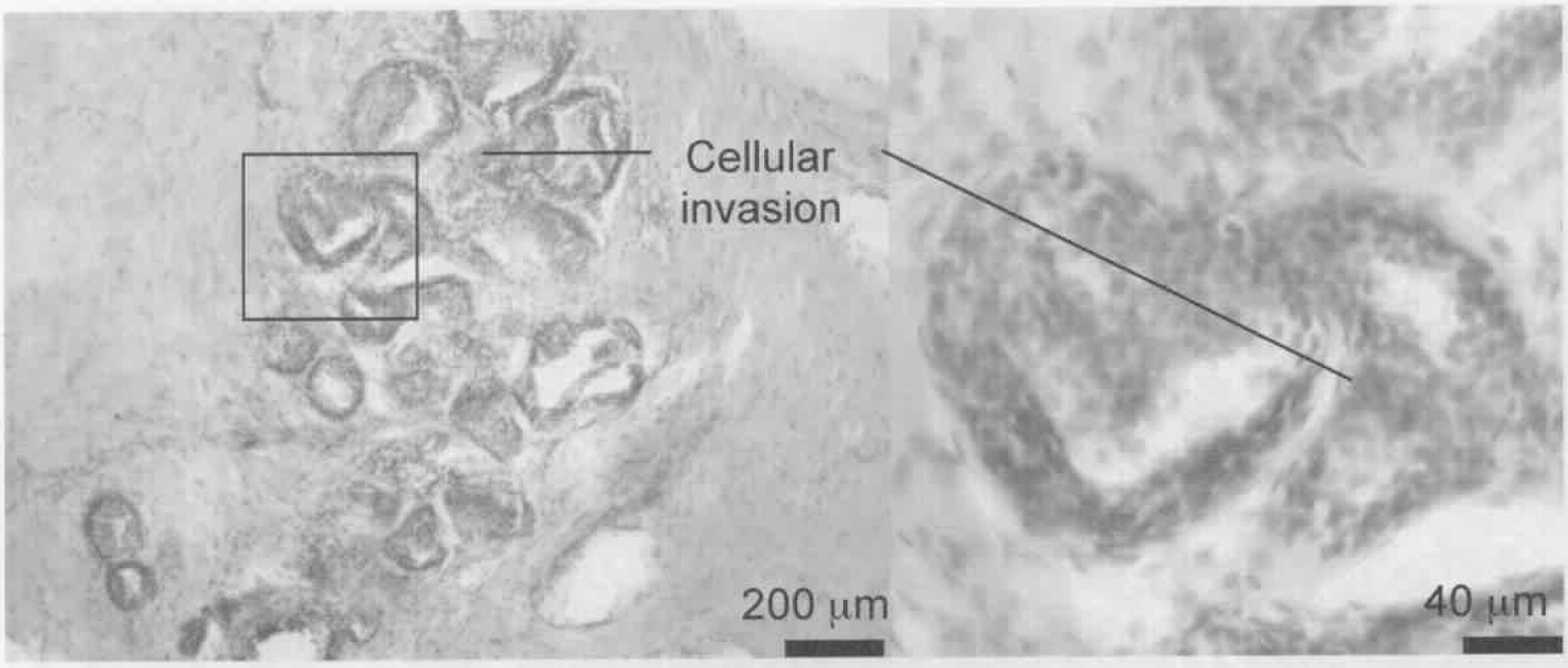


Figure 3.14. H&E of a FFPE section of an IDC, which displays multiple ducts with cellular invasion to the surrounding breast tissues. (Left) low magnification and (right) high magnification showing cellular invasion in multiple ducts.

3.5.2 INVASIVE LOBULAR CARCINOMA

Invasive lobular carcinoma (ILC) has histological characteristics that include nests of tumour cells (Figure 3.15) [2,4,6]. Here, large clusters of lobules are presented with cellular invasion. The tumour cells in the lobules of Figure 3.16 and the neighbouring duct display cellular infiltration to the surrounding tissues. The sizes of the nuclei are larger than those in normal cells.

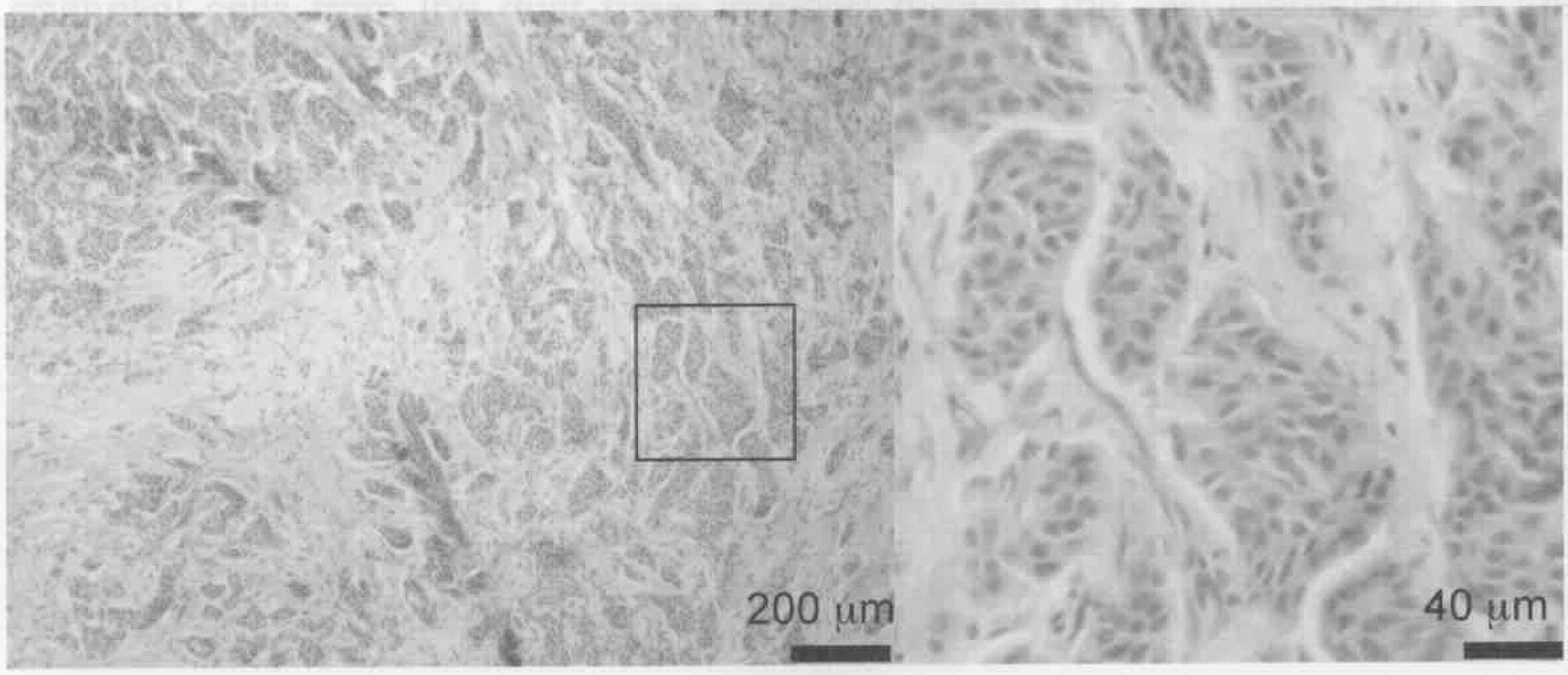


Figure 3.15. H&E of a FFPE section of an invasive lobular carcinoma showing multiple lobules filled with tumour cells. (Low) low magnification, (right) high magnification.

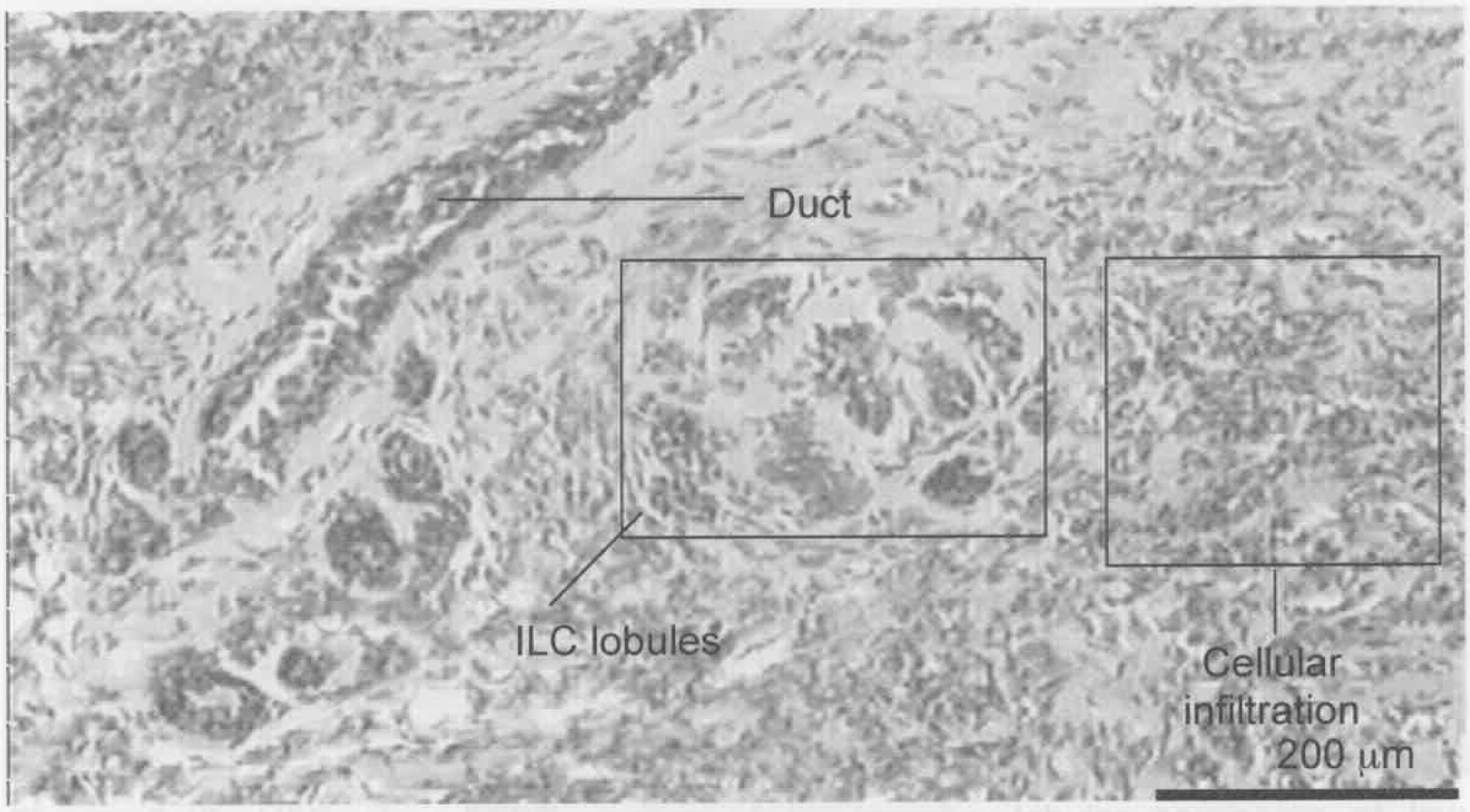


Figure 3.16. H&E of a FFPE section of an ILC showing the cellular infiltration around the lobules and the duct. Tumour cells are found in the surrounding connective tissues.

The ILC tumour cells can also present as a single sheet pattern as illustrated in Figure 3.17 [4]. The nuclei are small and uniform compared with the cellular infiltration in Figures 3.15 and 3.16. Figure 3.18 presents an ILC lesion displaying cancerous lobules, as well as the fibrocystic ducts with cellular infiltration. The lobules are clustered together with cellular infiltration while the epithelial cells from the duct have invaded through the cell membrane to the surrounding breast tissues.

Figure 3.17. H&E of a FFPE section of an invasive lobular carcinoma at low magnfication, which illustrates the typical, single-file infiltrative pattern of this tumour.

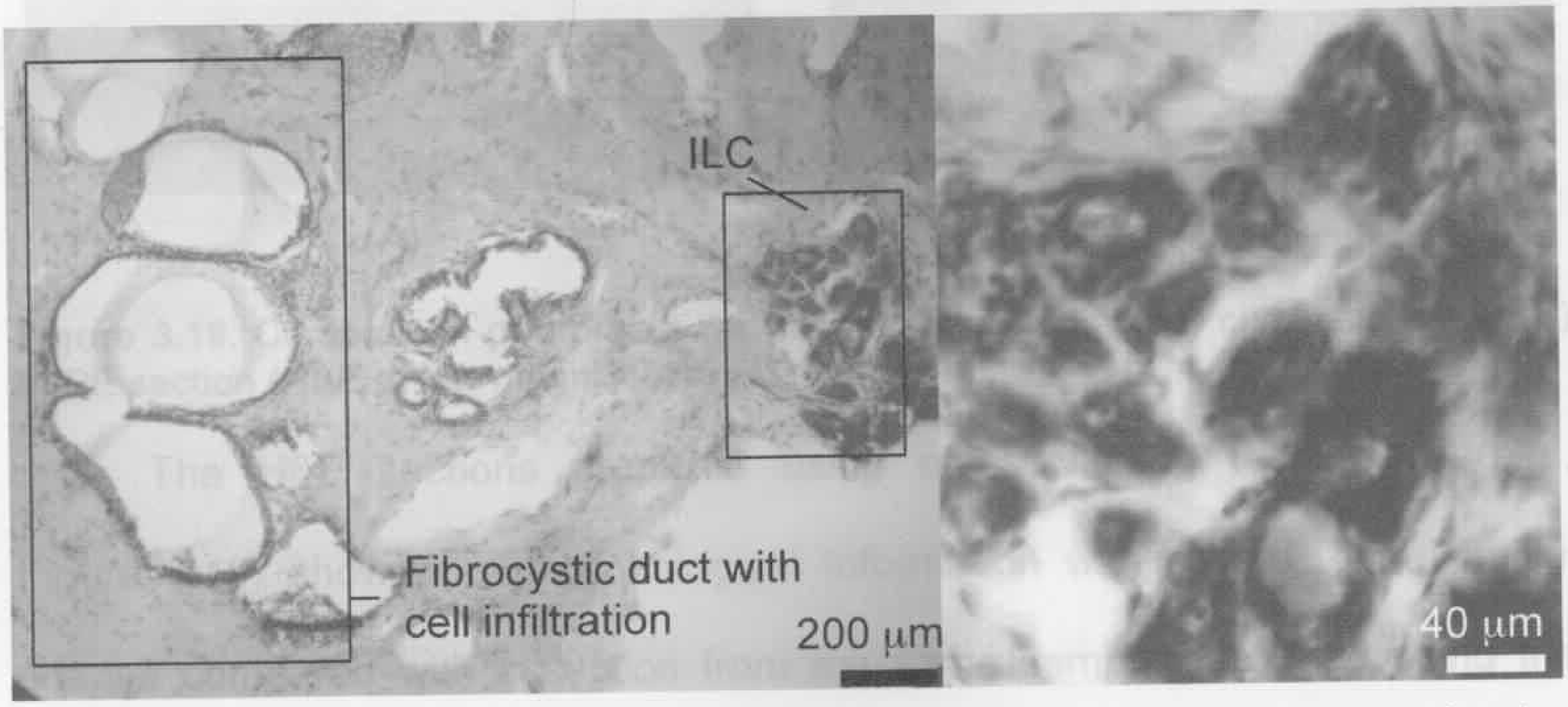


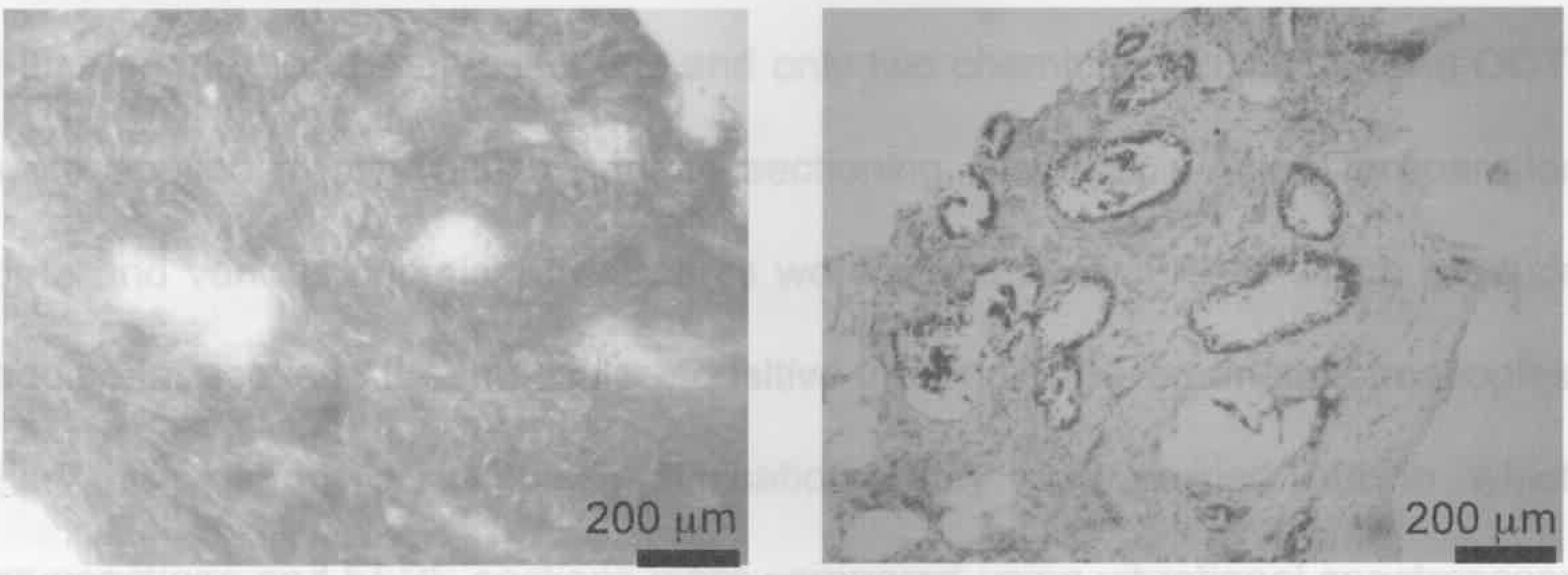
Figure 3.18. H&E of a FFPE section of an invasive lobular carcinoma showing infiltration of ducts by fibrocystic features with cellular invasion of a large duct and ILC on the right. (Right) high magnification of the ILC area.

3.6 DISCUSSION

Cryosections and FFPE sections were stained using H&E for diagnosis by a qualified pathologist. The preparation and preservation of these two section techniques were analysed (Table 3.2. and Figure 3.19) to compare with subsequent analyses by infrared and/or Raman spectroscopy on adjacent unstained sections.

Table 3.2. Comparison of cryosection and paraffin-embedded section.

	Cryosection	Paraffin-embedded Section
Preparation time	< 1 hr	Min. 24 hr
Chemicals used in preparation	2	Min. 8
Preservation time	Short	Long
Pathological features	Worse (if left too long)	Better



(A) (B)

Figure 3.19. Comparison of a breast IDC lesion presented as (A) cryosection and (B) FFPE section (H&E at low magnification).

The H&E sections prepared using cryosectioning techniques (e.g., Figure 3.19) show poor morphological information with regard to the nuclei shapes compared with a section from the same sample sectioned using the FFPE technique. The poor morphological features observed in the cryosection were apparent when the tissue was stored for an extended period (> 6 months). This feature is due to tissue degradation by autolysis as the section was not fixed in formalin, while there is no apparent cellular degradation seen on the FFPE section [1-4]. In addition, thin breast sections (< 8 μm) were difficult to prepare using the cryosectioning technique since high adipose content lesions required microtome operating temperatures between –30°C and –20°C compared to low adipose content lesion, which required a temperature between

–20°C and –15°C [1]. The sections curled prior to mounting on glass slides or infrared mounting substrates such that adjacent sections were not easily obtained using the cryosectioning technique.

The main criteria for section technique selection for the vibrational spectroscopic analysis was to minimise the interference from the chemicals used during sectioning processes. As seen in Table 3.2, the preparation time is much shorter for the cryosections and only two chemicals (PBS/D_2O and OCT) were applied to the tissues prior to sectioning, whereas a longer preparation time and various chemical treatments were required for FFPE, which produce additional peaks in the molecular-sensitive infrared and Raman spectroscopies. For this reason, a sample optimisation study was carried out in which cryosections and FFPE sections were compared using vibrational spectroscopy and the results are presented in Chapters Four and Six.

3.7 CONCLUSIONS

The histological features of the major breast diseases demonstrated why pathology is the current gold standard for disease identification. However, a more rapid and detailed diagnostic method would provide better follow-up treatments and increase the chance of survival. The major components of the tissues, diseases and malignancy were apparent in most sections. A comparison of two sectioning techniques, cryosectioning and FFPE sectioning, showed that FFPE sections retained all the tissue architecture whereas cryosections produced poor morphological features that may not be correctly identified by an experienced pathologist. In this study, three samples were unable to be classified due to a lack of morphological significance in the cryosections. By contrast, the minimal chemical treatment required for cryosectioning makes it the preferred sectioning technique for vibrational spectroscopic analyses.

3.8 REFERENCES

[1] J. Radford, *Histopathology Techniques.* 7th ed., The University of Sydney, Sydney, 2002.

[2] P. Rosen, *Rosen's Breast Pathology.* 2nd ed., Lippincott Williams & Wilkins, Philadelphia, 2001.

[3] S. Silverberg, *Atlas of Breast Pathology.* Elsevier Science (USA), Philadelphia, 2002.

[4] G. McKee, *Cytopathology of the Breast.* 1st ed., Oxford University Press, Boston, 2002.

[5] R. Fechner, S. Mills, *Breast Pathology: Benign Proliferations, Atypias and in-Situ Carcinoma.* 1st ed., American Society of Clinical Pathologists Press, Chicago, 1990.

[6] J. Harris, *Diseases of the Breast.* 3rd ed., Lippincott-Raven, Philadelphia, 2004.

[7] D. Innes, Intraduct Papilloma. http://www.med-ed.virginia.edu/courses/path/gyn/breast5.cfm, University of Vriginia Health System, Charlottesville, 2007.

[8] S. Komen, http://www.komen.org, The Susan G. Komen Breast Cancer Foundation (USA), Dallas, 2005.

[9] F. Greene, D. Page, I. Fleming, A. Fritz, C. Balch, D. Haller, M. Morrow, *Breast in American Joint Committee on Cancer: Ajcc Cancer Staging Manual.* 6th ed., Springer, New York, 2002, 223-240.

[10] H. Burstein, K. Polyak, J. Wong, S. Lester, C. Kaelin, *N. Engl. J. Med.* 350 (2004) 1430-1441.

[11] K. A. Skinner, *The Clinical Management of Ductal Carcinoma in-Situ, Lobular Carcinoma in-Situ and Atypical Hyperplasia of the Breast.* 1st ed., National Breast Cancer Centre, Sydney, 2003.

[12] K. A. Skinner, M. J. Silverstein, *Endocr. Rev.* 8 (2001) 33-45.

[13] E. Frykberg, *Breast J.* 5 (1999) 287-350.

4. SAMPLE OPTIMISATION FOR INFRARED SPECTROSCOPY

4.1 OPTIMISATION OF SAMPLE PREPARATION AND PRESENTATION

Initial studies were undertaken to establish and optimise the techniques used for sample preparation and presentation. The sampling techniques used were similar to standard sectioning procedures used in pathology laboratories. The areas of particular interest that were studied for optimising infrared spectroscopy diagnostics were: investigation of sectioning methods, development of methodologies to identify any unknown tissue presentation; evaluation of different infrared mounting substrates; and tissue heterogeneity issues faced in microspectroscopic investigations. The results of these experiments are presented in the following sections. An extensive tissue assignment of infrared bands can be found in Chapter 5, Section 5.1.

4.2 TISSUE FIXATION

Breast samples obtained from routine needle biopsies and/or surgical biopsies were stored in PBS/D_2O buffer solution (pH 7.0) then sectioned in two ways: cryosectioning or formalin-fixed paraffin-embedded (FFPE) sectioning. The detailed sample preparation processes are described in Chapter 2, Section 2.2. The chemicals used for tissue preparation (both pre- and post-sectioning) were examined by ATR spectroscopy to identify the presence of any exogenous chemicals within the spectra of processed tissues, which would complicate the diagnostic procedure (Figure 4.1).

The solution used for storage of the core biopsies was PBS/D_2O and its spectrum is shown in Figure 4.1A. The broad band centred at 3341 cm^{-1} is assigned to the ν(O–H) mode of the PBS solution [1]. The broad intense band centred at 2482 cm^{-1} is due to the ν(O–D) mode of the D_2O solution [1]. The band at 1240 cm^{-1} is assigned to the $\nu_{as}(PO_2^-)$ mode [1].

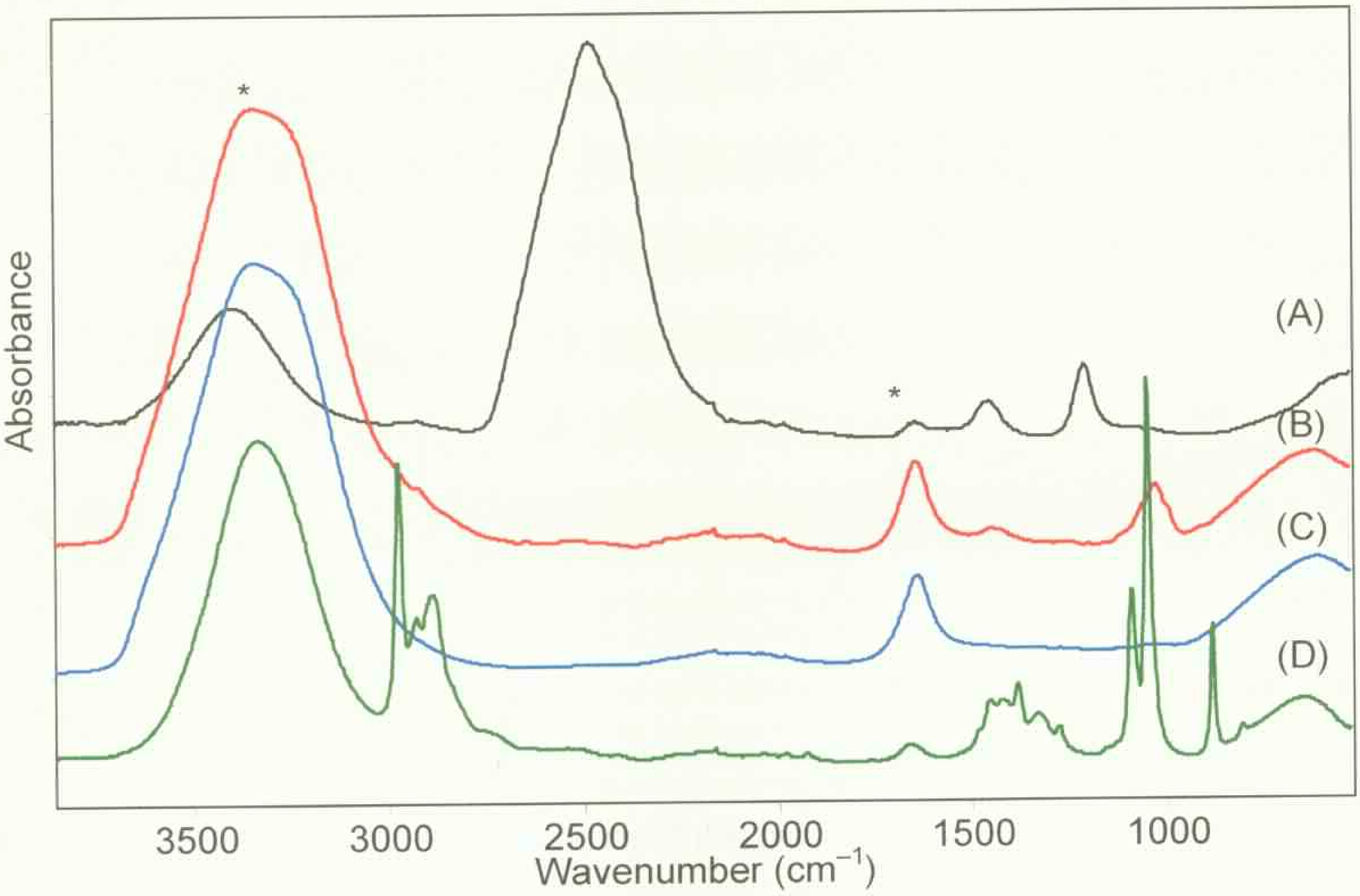


Figure 4.1. ATR spectra of chemicals used in tissue processing prior to embedding and sectioning: (A) PBS/D_2O; (B) 10% formalin solution; (C) tap water; and (D) 70% ethanol. Bands due to water vibrations are marked with asterisks (*). Experimental conditions: 256 scans; 4 cm^{-1}; and average of triplicates.

A standard formalin solution (10% v/v containing 4% formaldehyde) was used in tissue fixation to cross-link the proteins within the specimen, and prevent changes caused by autolysis and putrefaction [2-4]. The protein cross-linking also maintains the physical relationship between the cells and extracellular substances and makes the soluble substances within the cell insoluble, which assists in maintaining structural features for subsequent tissue processing [2-4]. Breast biopsies obtained for this study were typically small (<1 cm^3), and were fixed in formalin (4% formaldehyde, 10% v/v, pH 7.2–7.4) for a minimum of 12 hr for small lesions and 24 hr for larger samples in accord with standard procedures [3]. The ATR spectrum of the formalin solution in Figure 4.1B reveals a broad ν(O–H) band centred at 3280 cm^{-1}, which is attributed to water. The band at 1636 cm^{-1} is due to the aldehyde ν(C=O) and water δ(OH) modes and the 1030 cm^{-1} band is the aldehyde ρ(CH) mode [1].

After tissue fixation (at least 24 hr for maximum infiltration), the biopsy was washed with tap water to remove excess formalin solution. The spectrum (Figure 4.1C) of water is similar to that of formalin with broad features centred at 3280 and 1636 cm^{-1}, which are assigned to the ν(OH) and δ(OH) modes, respectively [1]. The washed tissue was subsequently stored in 70% ethanol (EtOH) prior to overnight tissue processing (Figure 4.1D).

An infrared spectrum of a breast biopsy stored in PBS/D_2O, formalin-fixed and subsequently cryosectioned is presented in Figure 4.2. The aim of this experiment was to evaluate whether the products used in formalin-fixation were evident in the infrared spectrum of the biopsy. The spectra of the formalin solution used to store the biopsy and a fresh PBS/D_2O solution are also shown. The spectral information obtained from the tissue is excellent: a small peak observed at ~2482 cm^{-1} is evidence of PBS/D_2O retained in the biopsy during H_2O/D_2O exchange prior to fixation. The peak at 1026 cm^{-1} is due to the presence of the formalin solution during tissue fixation. The protein amide I band is overlapped with the water δ(OH) band, which could affect the diagnostic information if the biopsies are measured in bulk. The adipose is retained as is evidenced by the peak at 1743 cm^{-1}, which is due to the ν(C=O) lipid ester band.

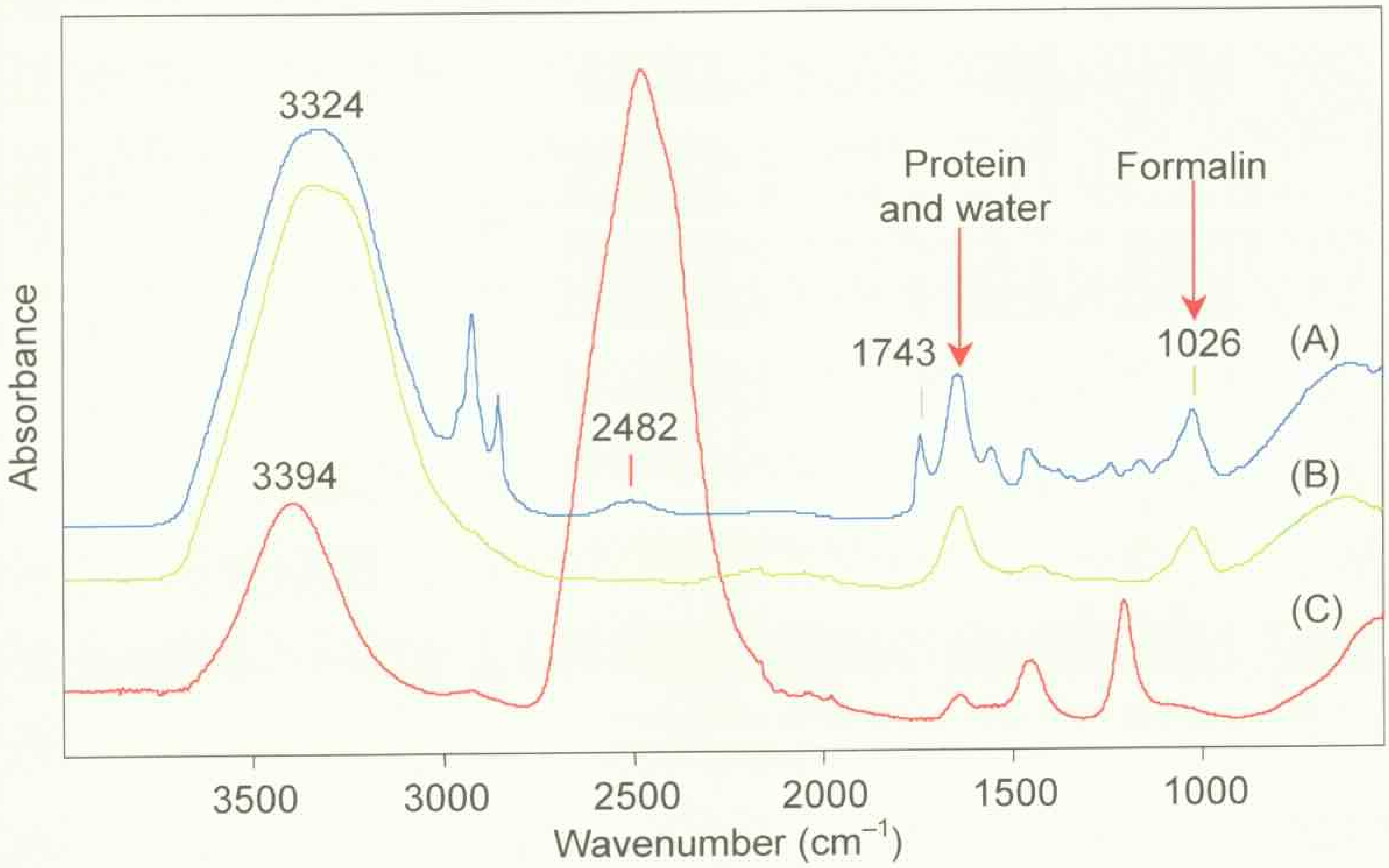


Figure 4.2. Significance of formalin-fixation in tissue sectioning: (A) a PBS/D_2O then formalin-fixed malignant breast biopsy; (B) formalin solution used in tissue fixation; and (C) PBS/D_2O solution used in the tissue storage. Experimental conditions: ATR spectroscopy; 512 scans; 4 cm^{-1}; purged with N_2; and average of triplicates. Spectra are to scale and are offset for clarity.

4.3 TISSUE PROCESSING

Once the tissues were fixed in formalin solution, they were processed overnight prior to FFPE sectioning. The tissue processing involved treatment with various grades of ethanol, paraffin wax and xylene solutions. The tissues were then mounted and oriented on plastic sample cassettes for paraffin-embedding, cooled, sectioned to the desired thickness, mounted on substrates (glass for pathology, or specific windows for vibrational spectroscopy), and dried overnight. They were deparaffinised in two changes of fresh deparaffinising agents (xylene or histopure) for 10 min each and stained using H&E according to the standard protocol outlined in Chapter 2, Section 2.2. ATR spectra of the chemical solutions used in tissue processing and staining are presented in Figure 4.3.

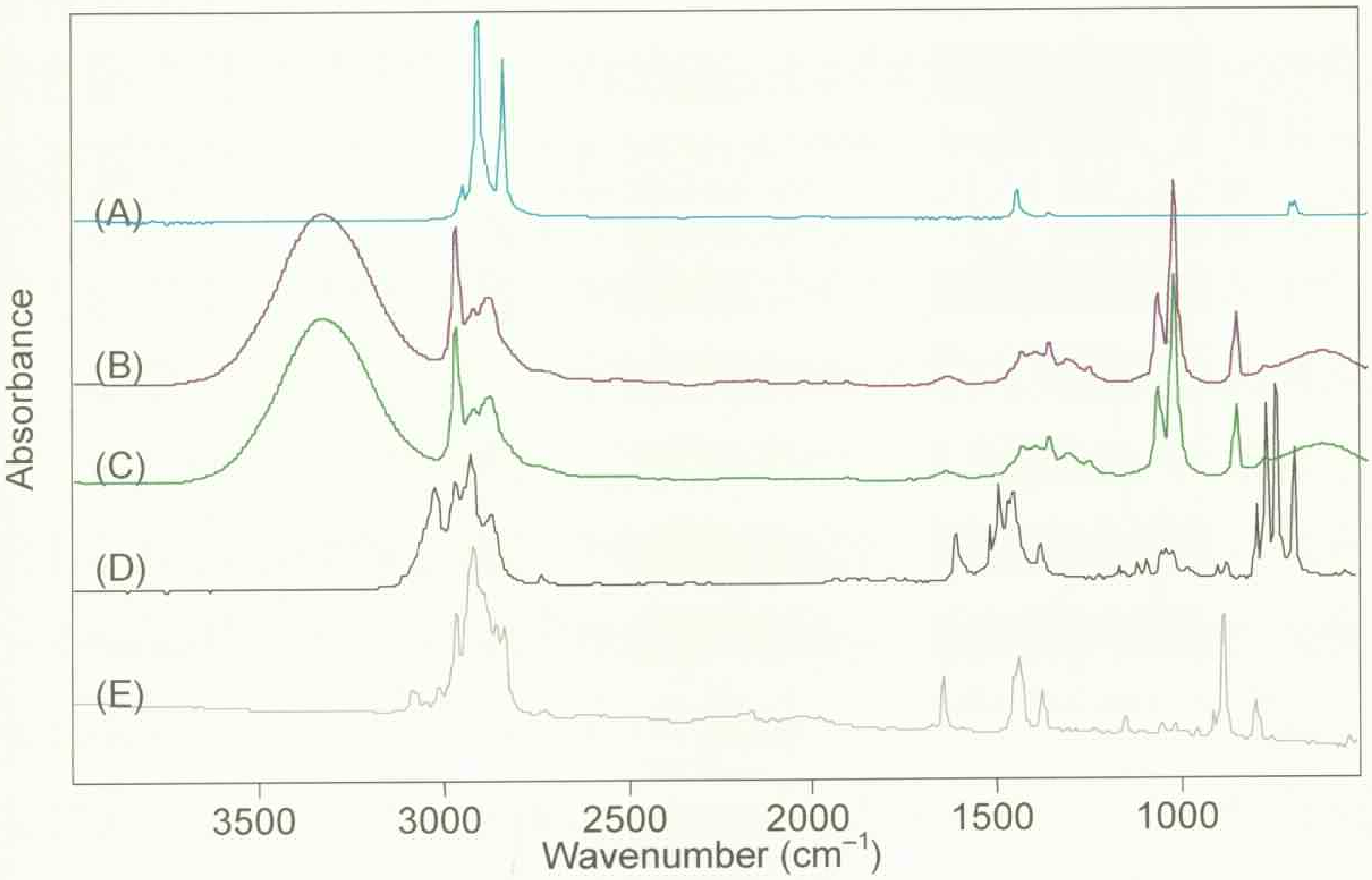


Figure 4.3. ATR spectra of chemical solutions used in tissue processing and staining process: (A) paraffin block; (B) EtOH 70%; (C) EtOH 95%; (D) xylene; and (E) histopure. ATR corrections were not applied to any of the spectra. Experimental conditions: 512 scans; 4 cm^{-1}; and average of triplicates. Spectra are offset for clarity.

In the short term, cryosectioning is the preferred technique, provided the sample is analysed within a week. An additional advantage of this technique is that the adipose content is retained. However, if samples are not fixed prior to cryosectioning, they will degrade due to autolysis, see the schematic workflow and sample processing chart in Figure 2.1. Paraffin embedding allows the tissues to be fixed and preserved in a way that allow retrospective analysis [3]. The spectrum of the paraffin block (Figure 4.3A) is dominated by strong $\nu_{as}(CH_2)$ and $\nu_s(CH_2)$ absorptions at 2914 and 2846 cm^{-1}, $\delta(CH_2)$ modes at 1471 and 1462 cm^{-1}, and C–H rocking modes at 728 and 720 cm^{-1} [1].

As expected, the two concentrations of ethanol (70% and 95%) have similar spectra and bands are due to the OH, CH_2 and CH_3 groups (Figures 4.3B and 4.13C). Xylene and histopure solutions are used as deparaffinising agents during tissue processing and prior to staining. The infrared spectra of xylene and histopure (Figures 4.3D and 4.13E, respectively) have vibrations due to C=C, C=O and CH modes of the organic structures.

4.4 TISSUE SECTIONING

IR spectra from tissue sections prepared using cryosectioning and FFPE (both prior to and post-deparaffinisation) sectioning techniques are shown in Figure 4.4. The spectrum of a cryosection (Figure 4.4A) is typical of a biological tissue sample containing adipose tissue (arrows), whereas both paraffinised FFPE (Figure 4.4B) and deparaffinised (Figure 4.4C) sections displayed traces of chemicals such as paraffin wax used during FFPE processing (marked with asterisks *). Although the deparaffinised section (Figure 4.4C) was immersed in two changes of xylene/histopure solutions for 10 mins each to remove paraffin according to the standard pathology protocol [3], clearly the process did not remove the paraffin completely. Prolonged treatment of tissue with deparaffinising agents is not recommended as it could cause tissue shrinkage, particularly in fibrous tissue, therefore, the absolute minimum times (10 min each) for deparaffinisation should be used [3].

Previous research by Liu and co-workers observed two bands at 1474 and 1464 cm^{-1} in the spectra of breast sections using synchrotron-radiation based Fourier transform infrared (SR-FTIR) spectroscopy that have not been previously reported when using a conventional IR spectrometer with a globar source [5]. These two bands were found to have higher intensities in malignant tumours than in normal and benign tissues. However, in this current study and from the literature, these bands have been assigned to $\delta(CH_2)$ bands, which originate from the long hydrocarbon-chain of paraffin wax (Figure 4.5) [6,7].

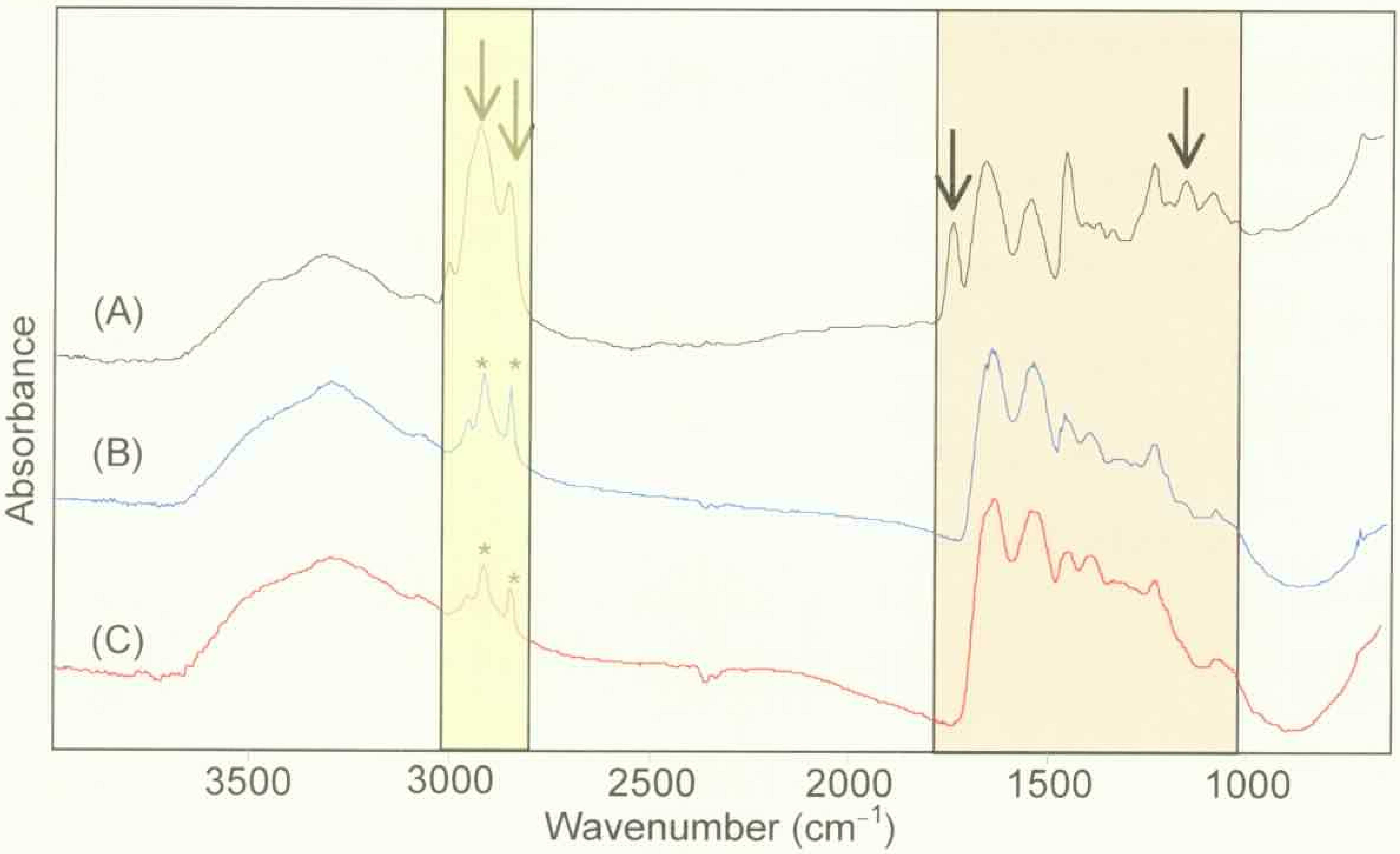


Figure 4.4. IR spectra of a benign breast tissue sample analysed as: (A) a cryosection; (B) a paraffinised FFPE section; and (C) a deparaffinised FFPE section after two changes of fresh xylene and histopure solutions (10 min each). Peaks are due to adipose tissues are marked with arrows and those due to paraffin wax are marked with (*). Experimental conditions: 5-µm-thick sections mounted on transflective slides; 512 scans; 4 cm^{-1}; and ×15 magnification. Spectra are offset for clarity.

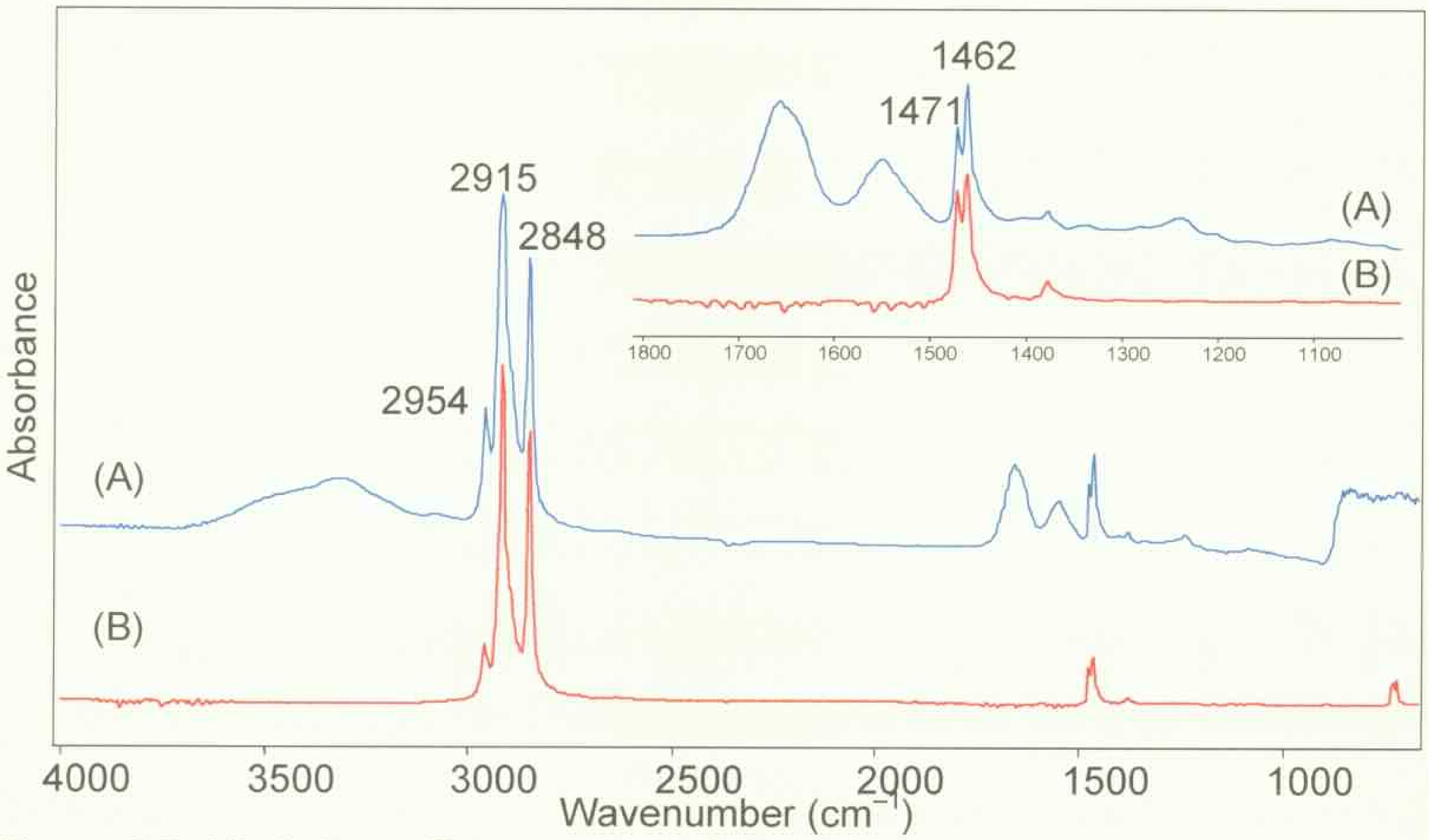


Figure 4.5. (A) A deparaffinised malignant breast FFPE section exhibiting residual paraffin not removed by deparaffinising solutions. (B) IR spectrum of paraffin from the paraffin block (storage) of the same biopsy. Inset: Expanded region of 1800–1000 cm^{-1} showing the two paraffin bands at 1471 and 1462 cm^{-1} that interfere with the spectral characterisation. Experimental conditions: FTIR microspectroscopy; 512 scans; 4 cm^{-1}; and tissue was mounted on transflective slide for transflection experiment at ×15 magnification. Spectra are offset for clarity.

4.5 DETERMINATION OF TISSUE SECTIONING TYPES USING PRINCIPAL COMPONENT ANALYSIS

One of the advantages of using IR spectroscopy with multivariate statistical analysis, such as principal component analysis (PCA), is the ability to identify and separate subtle variations in biological spectra and to identify residual chemicals remaining from storage or processing. As seen in Figure 4.3, these exogenous chemical solutions have prominent ν(OH) and ν(NH) bands in the 3600–3050 cm^{-1} region and ν(CH) bands within the 3050–2750 cm^{-1} region. Therefore, careful examination of this region could potentially identify any tissue contaminants. IR spectra were collected (not shown) from formalin-fixed cryosections and paraffinised FFPE sections of three malignant biopsies together with the PBS/D_2O and formalin solutions used for storage and fixation, respectively. A PCA analysis of the 3050–2750 cm^{-1} region and the resulting PC2 *vs.* PC1 plot is displayed in Figure 4.6.

PC1 separates the formalin-fixed cryosections and FFPE sections, which accounted for 99.7% of the spectral variance, while PC2 (0.3%) separates the formalin solution from the formalin-fixed cryosections. The previous section demonstrated that the strong $\nu_{as}(CH_2)$ and $\nu_s(CH_2)$ absorptions in the 3050–2750 cm^{-1} region observed in FFPE sections are due to paraffin contamination. From this preliminary analysis, the ν(CH) region can be used to distinguish the tissue sectioning method, and also demonstrates that any diagnostic procedure will need to take into account the methods used for tissue preparation.

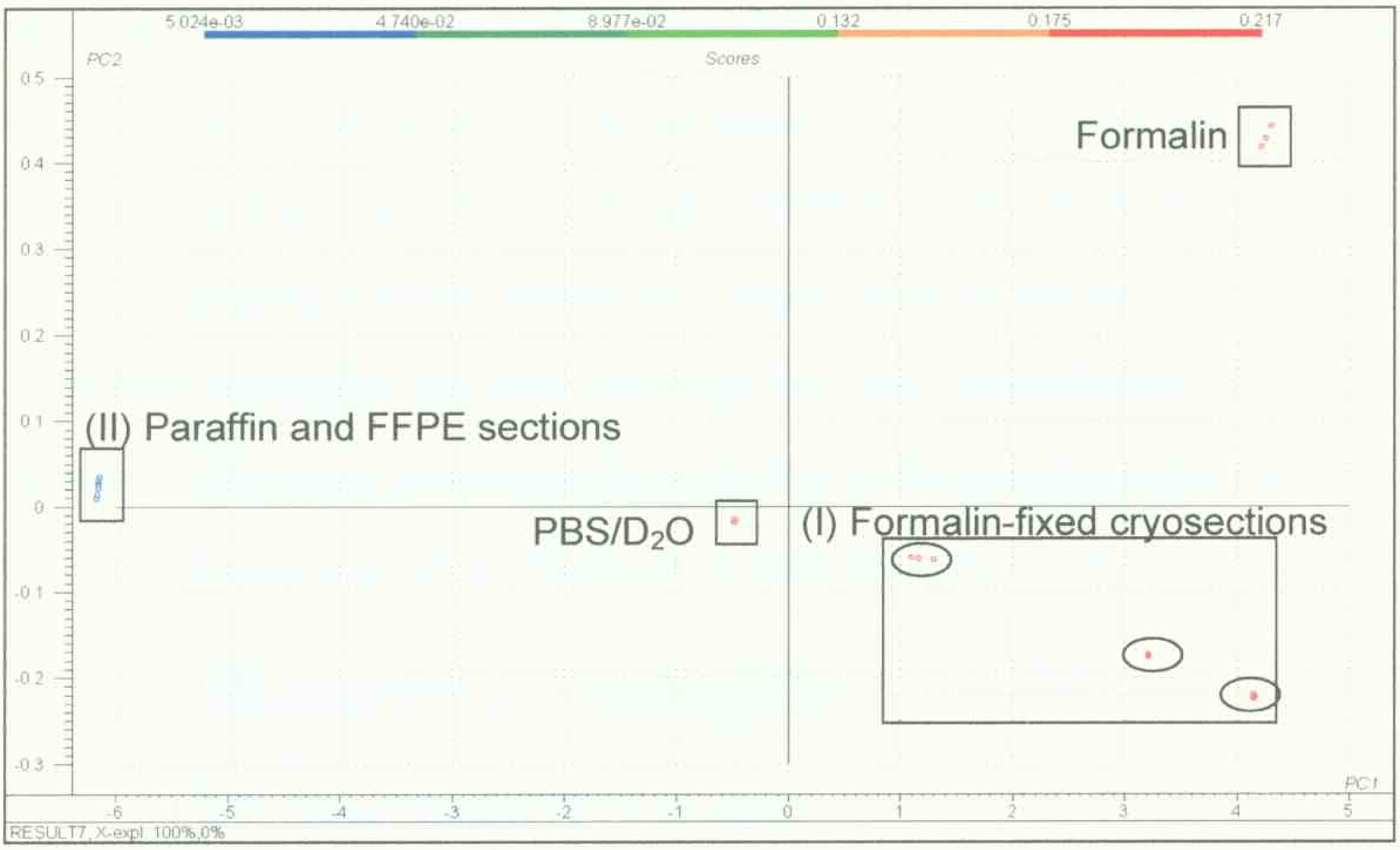


Figure 4.6. PC2 *vs.* PC1 scores plot of a training data set containing three malignant breast biopsies prepared in: (I) formalin then cryosectioned; and (II) in paraffin then FFPE sectioned, and the storage solutions of PBS/D_2O and formalin, in the spectral region of 3000–2750 cm^{-1}. Experimental conditions: 512 scans; 4 cm^{-1}; and ×15 magnification; and each point represents an average of triplicate measuring at the cancerous regions.[1]

Figure 4.7 compares the various tissue sectioning techniques using PCA on one malignant IDC breast biopsy using the spectral region of 3050–2750 cm^{-1}. The use of one biopsy was to minimise sample variations, i.e., the PCA experiment was designed to examine only the variations due to sectioning techniques. Each point on the PC plot represents an average of three spectra collected at the tumour cells within the ducts. Paraffinised FFPE sections are separated from the deparaffinised FFPE sections in PC1 (74% of total variance), while tissues that were cryosectioned directly after storage in PBS/D_2O are separated from those of formalin-fixed cryosections. PC2 (14% of total variance) describes the separation of the cryosections and FFPE sections,

---

[1] Only integer values were displayed in The Unscrambler PC score plots, PC1 and PC2 spectral variance should be 99.7% and 0.3%.

although there is a closer relationship of formalin-fixed cryosections to the FFPE sections. The separation can be explained in terms of the stages of the tissue processing, stage (I) being storage of PBS/D_2O then cryosectioned or the biopsy undergoing formalin fixation (II). Some sections will then go on to be embedded in paraffin and then sectioned (III) and deparaffinised (IV). The variations within each processing stage are due to the heterogeneous nature of the breast tissues and will be discussed in detail in Section 4.10.

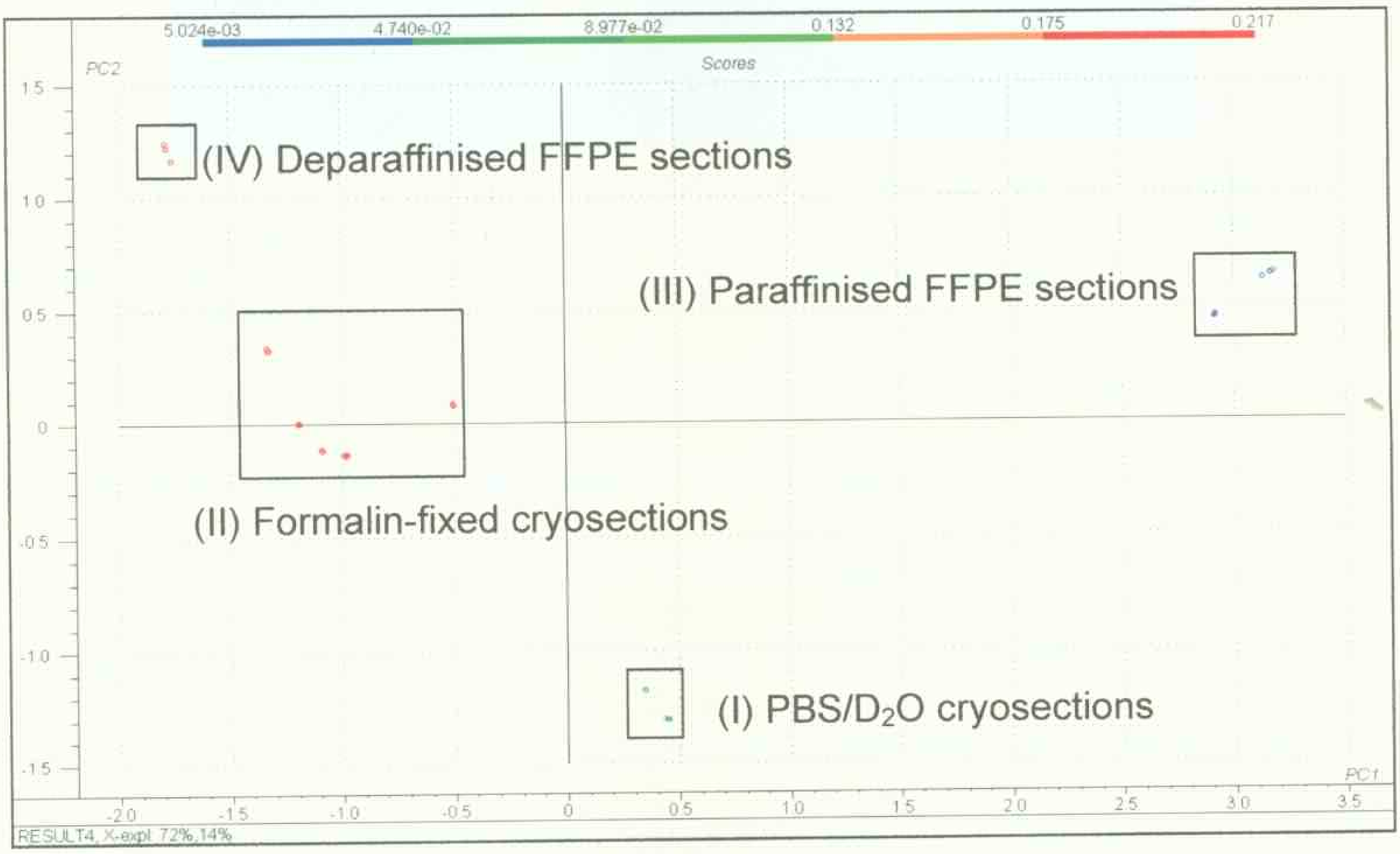


Figure 4.7. PC2 *vs.* PC1 scores plot separating tissue presentation as (I) PBS/D_2O cryosections, (II) formalin-fixed cryosections, (III) FFPE sections; and (IV) deparaffinised FFPE sections in the spectral region of 3050–2750 cm^{-1}. Sections from one malignant IDC breast biopsy were used here. Experimental conditions: 512 scans; 4 cm^{-1}; and ×15 magnification. The variations seen within each groups are due to the heterogeneity nature of the breast tissues where the IR spectra were collected.

4.6 TISSUE STABILITY

Formalin-fixed biopsies and subsequent FFPE sections were stored at room temperature for three and half years without any apparent signs of tissue degradation (Figure 4.8A), whereas samples that were stored in PBS/D_2O and not fixed were found to have deteriorated after three months due to autolysis, and could not be retrospectively analysed (Figure 4.8B).

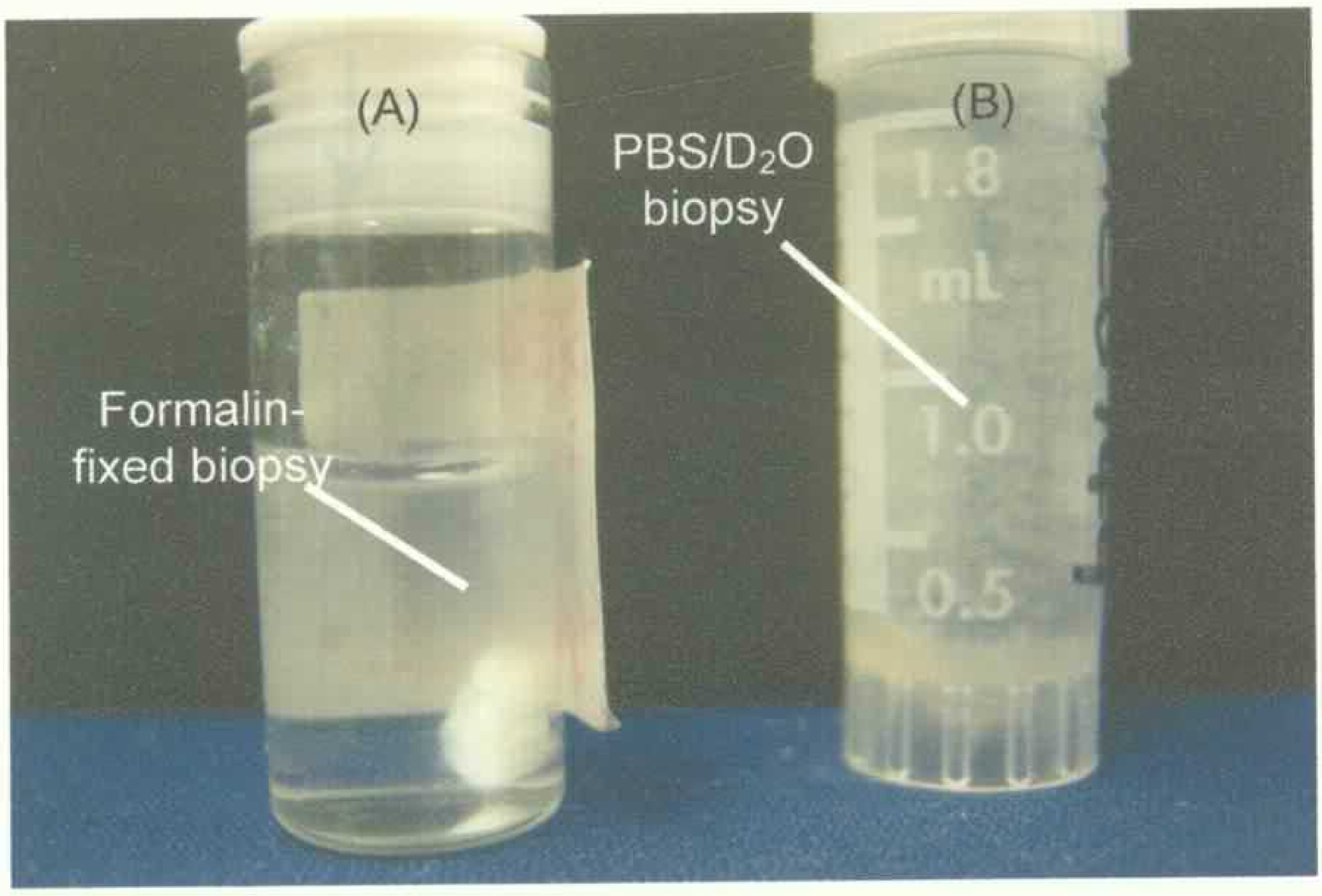


Figure 4.8. Comparison of tissue biopsy after three and half years of storage. (A) A formalin-fixed biopsy whose tissue structure has been preserved and (B) a deteriorated breast biopsy stored in PBS/D_2O, the biopsy has disintegrated and is suspended throughout the solution.

Although the formalin-fixed biopsies do not appear to have deteriorated visually this may not be the case at a molecular level. Therefore, infrared spectroscopy was used to establish if tissue degradation had occurred in two formalin-fixed malignant breast biopsies stored for three and half years. Initially, these samples were prepared as cryosections and were examined using IR microspectroscopy (see Figures 4.9A and B). When re-evaluated, the unprocessed biopsies were examined using ATR spectroscopy (Figures 4.9C and D). There was insufficient sample volume remaining to prepare a cryosection, therefore the sample was measured in bulk using ATR spectroscopy.

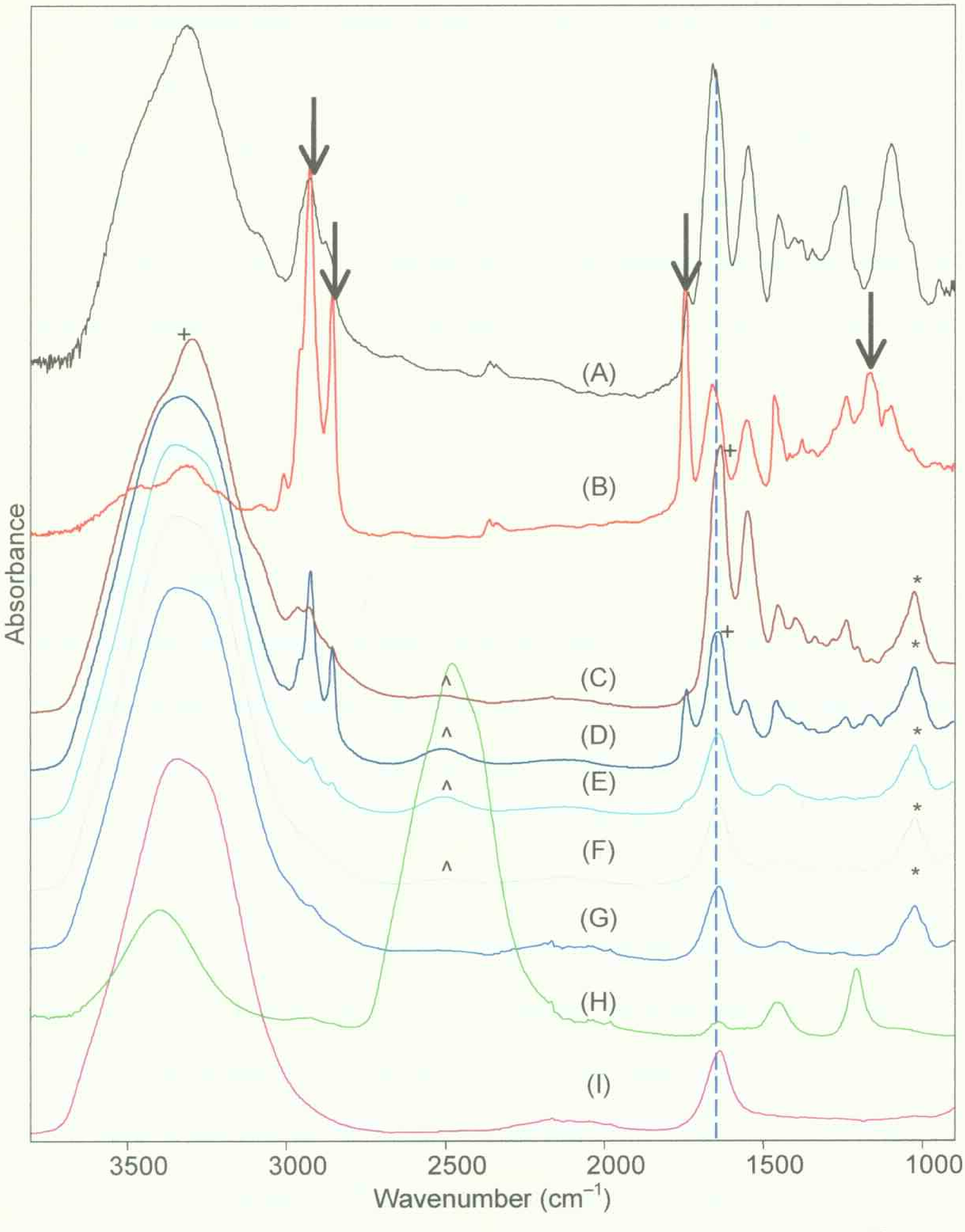


Figure 4.9. Spectral evaluations of the effects of tissue fixation three and half years after sample acquisitions. (A & B) Original IR spectra of two malignant cryosections measured microscopically; ATR spectra of (C & D) two three and half year-old, formalin-fixed malignant breast specimens measured in bulk; (E & F) formalin solutions used in fixation and storage in samples A and B, respectively; (G) fresh formalin solution, (H) fresh PBS/D_2O and (I) tap water. Experimental conditions: (A and B) microscopy: 512 scans; 4 cm^{-1}; and ×15 magnification. (C–I) ATR spectroscopy: purged with N_2; 128 scans; and 4 cm^{-1}. Peaks that are attributed to adipose tissues are marked with arrows; PBS/D_2O with (^); formalin solution with (*) and water with (+). The amide I band is highlighted with blue line. Spectra are offset for clarity.

The biopsies were initially stored in PBS/D_2O solution, then subsequently fixed in formalin solution and stored at room temperature. The spectrum of the storage solution together with spectra of fresh formalin solution, PBS/D_2O, and tap water are shown in Figures 4.9E–I. The assignment of the vibrational bands observed in the cryosectioned and biopsy spectra will be discussed in detail in Chapter 5, Section 5.1. The objective of this particular experiment was to establish if the molecular structure of the tissue had deteriorated with time and the discussion will be limited to the presence of any such indicators.

Small contributions from the ν(O–D) bands of the PBS/D_2O solution are observed at ~2500 cm^{-1} in both biopsies, which were originally stored in this solution prior to formalin fixation (marked with (^), Figures 4.9C and D). Therefore, there were some H_2O/D_2O exchanges within the biopsies. The peaks marked with (*) and (+) are due to the infrared absorptions of the formalin solution and water. The IR spectrum of the malignant tissue biopsy in Figure 4.9D displays stronger bands of $\nu(CH_2)$ at 2921 and 2852 cm^{-1}, $\nu(C{=}O)_{\text{lipid-ester}}$ at 1743 cm^{-1} and δ(C–C) at 1163 cm^{-1} that are attributed to the adipose tissues (see arrows). No bands due to adipose tissues were apparent in Figure 4.9C but strong bands were observed due to the protein absorptions. The aldehyde ρ(CH) band of formalin at 1030 cm^{-1} is observed in the biopsy samples (Figures 4.8C and D) but is absent in the spectra of cryosections (Figures 4.9A and B).

The amide I band shifts from 1655 cm^{-1} in the spectra of cryosections (Figures 4.9A and B) to 1640 cm^{-1} in the formalin-fixed biopsies (Figures 4.9C and D). The shift is due to the protein cross-linking during formalin-fixation as reported in the literature [3,6,8]. No signs of protein degradation in formalin-fixed biopsies are observed in this experiment. The IR spectra of the original cryosections were both stored in PBS/D_2O solution and cryosectioned without formalin-fixation, which preserved the adipose content as indicative by the lipid bands. As seen here, formalin-fixation does not affect any adipose content of the tissues and could be used for retrospective analysis.

Weak IR bands from the PBS/D_2O solution were observed (Figures 4.9E and F) only in the formalin solutions used for storing the biopsies. This effect was not observed in fresh formalin solutions (Figure 4.9G) since D_2O/H_2O exchanged between the biopsy and the formalin solution occurs over time. Although the water δ(OH) band at 1650 cm^{-1} (Figure 4.9I) is of medium intensity, the peaks observed in both original cryosections (Figures 4.9A and B) and formalin-fixed biopsies (Figures 4.9C and D) are due to the protein ν(C=O) bands from the biological tissues. Both cryosections were air-dried piror to spectroscopic analysis.

4.7 TISSUE THICKNESS

Tissues sectioned for histopathology have thicknesses of between 5 and 10 μm, which do not alter tissue architecture compared to thicker sections (>20 μm) [2-4,9-12]. In order to establish the optimal thickness for IR spectroscopic analysis, a LCIS biopsy was cryosectioned at various thicknesses (5, 10, 25, 40 and 50 μm) and mounted on CaF_2 windows. As shown in Figure 4.10, the IR spectra of all tissue sections are very similar, with the most significant difference being differing levels of adipose content.

The 5-μm and 50-μm-thick breast sections were difficult to use as they curled prior to mounting. In addition, the spectra of thicker samples are more likely to contain a number of overlapping tissue components (e.g., ducts and lobules), therefore confounding spectral analysis (Figure 4.10C). Ultimately, the goal is to produce a methodology that can be used in parallel with histopathology, which typically use tissue sections with thicknesses of between 5 and 10 μm. As the IR spectra of sections of these particular thicknesses are of good quality and give high signal-to-noise (S/N) ratio, it was determined that they were the optimal thicknesses for further spectroscopic analysis.

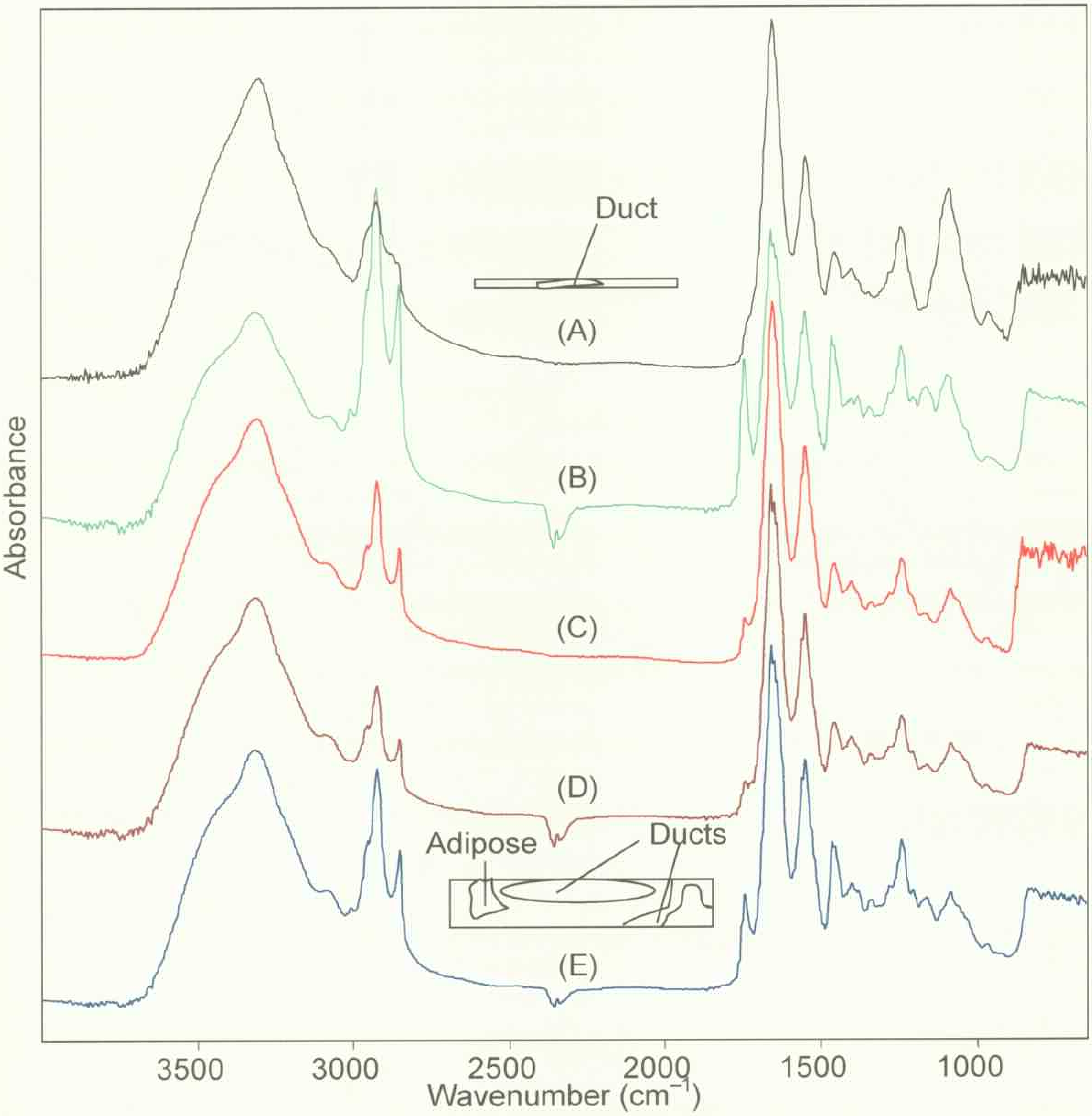


Figure 4.10. Comparison of a lobular carcinoma in-situ breast biopsy cryosectioned at various thicknesses: (A) 5 μm; (B) 10 μm; (C) 25 μm; (D) 40 μm; and (E) 50 μm. Experimental conditions: 512 scans; 4 cm^{-1}; ×15 magnification; and sections were mounted on CaF_2 windows. Spectra are to scale and offset for clarity. Boxes are schematic diagrams representing morphological components of tissue sections with thicknesses of 5 μm and 50 μm.

4.8 MOUNTING SUBSTRATES

Samples of an optimal 5–10 μm thickness were mounted on three IR mounting substrates, a CaF_2 infrared window, a transmission-reflective (transflective) slide and a standard silver-based mirror. The results (Figure 4.11) show that the CaF_2 window produces a spectrum with the flattest baseline, whereas baseline correction procedures are required for sections mounted on a transflective slide or a mirror. No signals are observed below 900 cm^{-1} (Region II) in the spectra of the tissues mounted on CaF_2 and below 600 cm^{-1} in the spectra on the transflective slides as both materials absorb all radiation below these regions. Similarly no signal is observed below 900 cm^{-1} in the cryosection mounted on mirror.

The spectrum of the tissue on the transflective slide (Figure 4.11 B) has an artefact within the 2300–1700 cm^{-1} spectral region (Region I). The artefact interferes the lipid band near 1750 cm^{-1} and the amide I band. This spectral artefact is more apparent near sample edges and at duct openings, and is ascribed to the anomalous dispersion artefact. Romeo and Diem have reported that if the artefact is not corrected, it can prevent the correct differentiation of tissue types in a sample [13,14]. This is especially important in the interpretation of IR imaging or mapping data with multivariate statistical analysis.

Calcium fluoride as a substrate will produce a spectrum with a flat baseline and good S/N. The window is also resistant to aqueous and organic solvents; however, the slide is more expensive than the transflective Ag/SnO_2 slide and mirror. The CaF_2 windows and mirror were re-used, but once sections were removed, they could not be re-examined. Transflective slides are a more attractive alternative because they are relatively inexpensive, the sections can be mounted on the slides permanently and can be re-examined at a later time with comparable spectral quality to those sections mounted on CaF_2 windows. In addition, the tissues can be stained after spectroscopic analysis and sent to histopathology for examination.

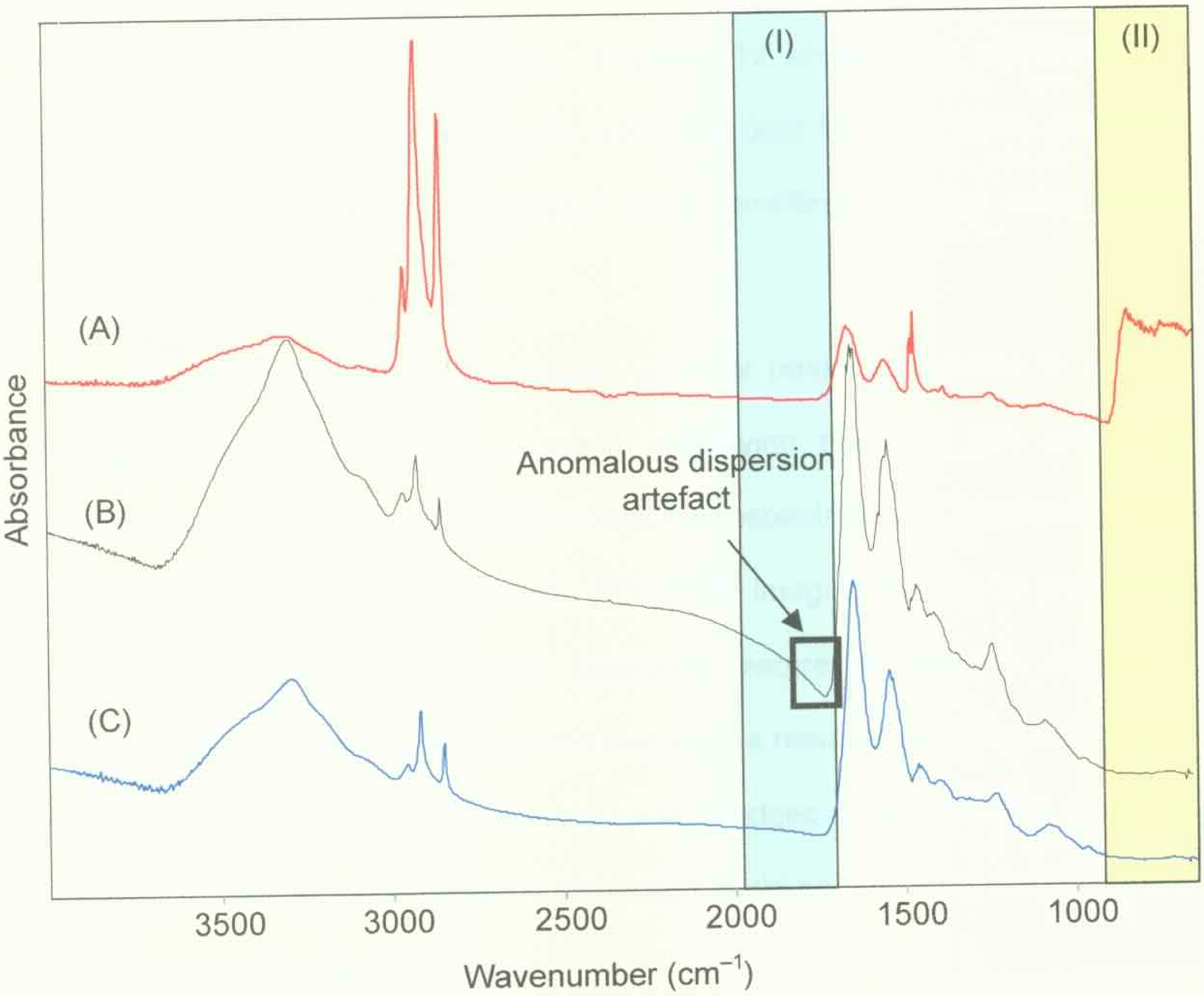


Figure 4.11. Comparison of IR spectra obtained from tissue sections mounted on various tissue-mounting materials: (A) FFPE section on CaF_2 window; (B) FFPE on transflective slide; and (C) FFPE section on mirror. Experimental conditions: all samples were 5 μm-thick; 512 scans; 4 cm^{-1}; and ×15 magnification. Spectra are offset for clarity.

4.9 SPECTRAL COMPARISON OF SINGLE-POINT, FPA-FTIR IMAGING AND SR-FTIR MAPPING

Different sampling processing and mounting substrates have been shown to yield different spectral results. Similarly, subtle spectral variations were also observed using different spectroscopic techniques: single-point microspectroscopy, focal-plane array infrared imaging, and synchrotron-radiation source infrared mapping. Figure 4.12 shows four representative spectra collected at the tumour cells within the duct from the same cribriform DCIS lesion (FFPE section mounted on a transflective slide) using these techniques.

The spectra are similar with all the major peaks assignable to various biological components. The spectrum displaying the best S/N and least artefacts was collected using single-point microspectroscopy. Following this, the spectrum was collected with the FPA-FTIR imaging system to produce a spectrum that displays a strong dispersive feature as seen in Region I. However, as previously discussed, the feature is a result of spectra having been collected from areas that are near the sample edges or around duct openings [13,14]. Therefore, the anomalous dispersion could be, and was, also observed in single-point microspectroscopy.

Lastly, two DCIS spectra collected under the same experimental conditions (32 scans; 4 cm^{-1} spec. res.; aperture size of 15×15 μm^2) were compared (Figures 4.12 C and D) using SR-FTIR microspectroscopy, but were recorded using different mirror velocities. Although the SR-FTIR spectrum would be expected to have a very high S/N ratio, Figure 4.12 C illustrates that significant noise was observed, particularly within Region II. This observation was attributed to the slower modulation frequency (i.e., velocity of moving mirror) that was chosen (0.4747 cm/s in Figure 4.12 C and 1.8988 cm/s in Figure 4.12 D). The S/N in Figure 4.12 D was improved when faster modulation frequency was used, and the $\nu_{as}(PO_2^-)$ of nucleic acid band at 1240 cm^{-1} was observed.

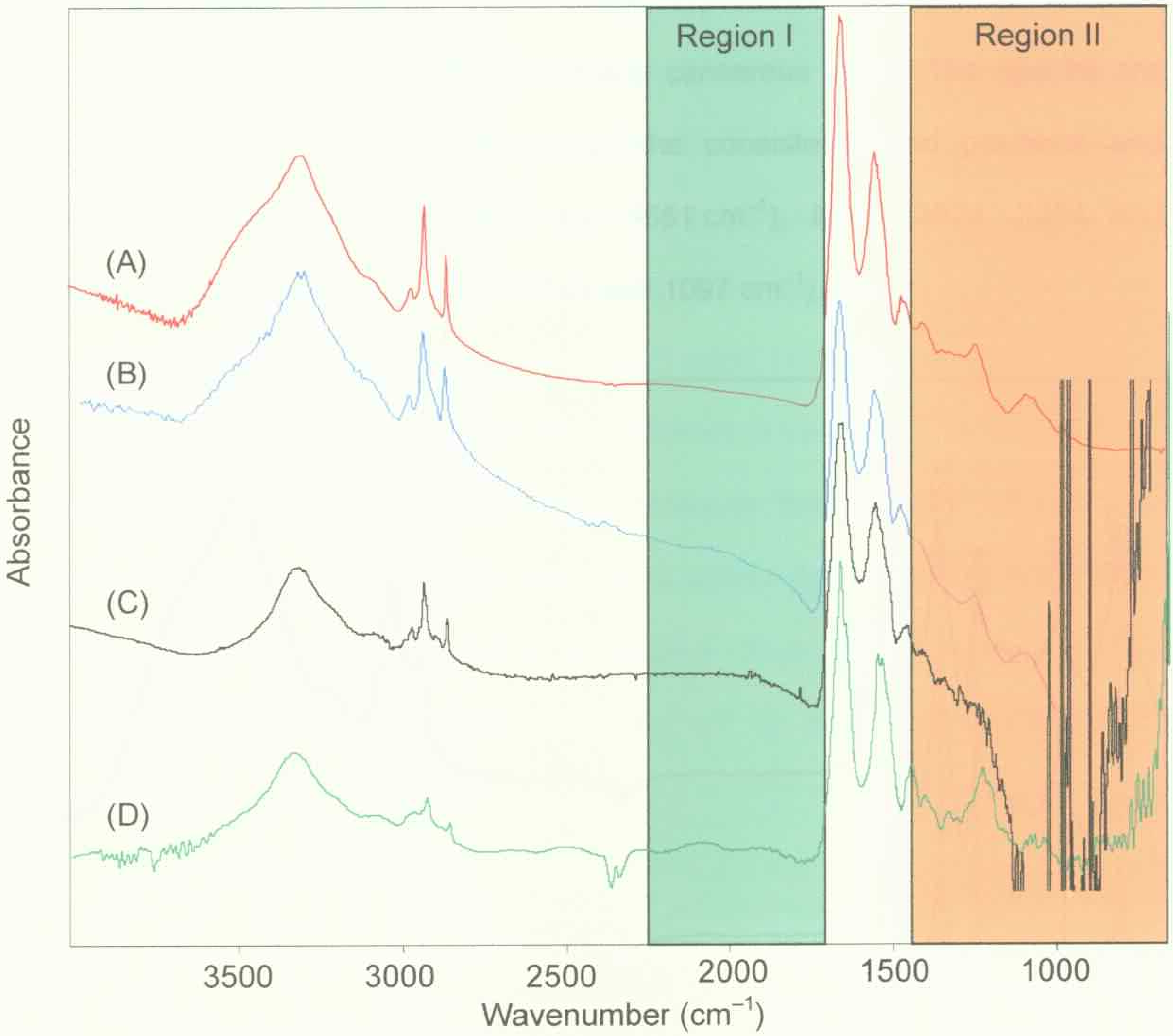


Figure 4.12. IR spectra of a low-grade cribriform DCIS FFPE section mounted on transflective slide using: (A) single-point microspectroscopy (512 scans; 4 cm^{-1}; ×15 magnification and aperture size of 5×5 mm^2); (B) FPA-FTIR imaging (64×64 array; 256 scans; 2 cm^{-1}; ×15 magnification; and aperture size of 5×5 mm^2); (C) SR-FTIR mapping (32 scans; 4 cm^{-1}; mirror velocity 0.4747 cm/s; ×32 magnification; and aperture size of 15×15 μm^2); and (D) SR-FTIR mapping (32 scans; 4 cm^{-1}; mirror velocity 1.8989 cm/s; ×32 magnification; and aperture size of 15 ×15 μm^2). Region I: dispersive artefact affected regions and Region II: SR-FTIR affected by modulation frequency and diffraction limit. Spectra are not to scale and are offset for clarity.

4.10 SPECTRAL REPEATABILITY AND TISSUE HETEROGENEITY

Numerous single-point spectra from samples of the same classification and from various positions on the same tissue type (i.e., lobules, ducts, adipose and connective tissues) were obtained to establish the reproducibility and repeatability of the analytical technique. Figure 4.13 shows three spectra of from an invasive ductal carcinoma breast cryosection measured at different positions of tumour cells within the same cancerous duct. The spectra are highly repeatable as is evidenced by the consistent band positions and lineshape of the proteins (1659 and 1551 cm^{-1}), lipids (2924, 2854 and 1454 cm^{-1}) and the nucleic acids (1241 and 1097 cm^{-1}).

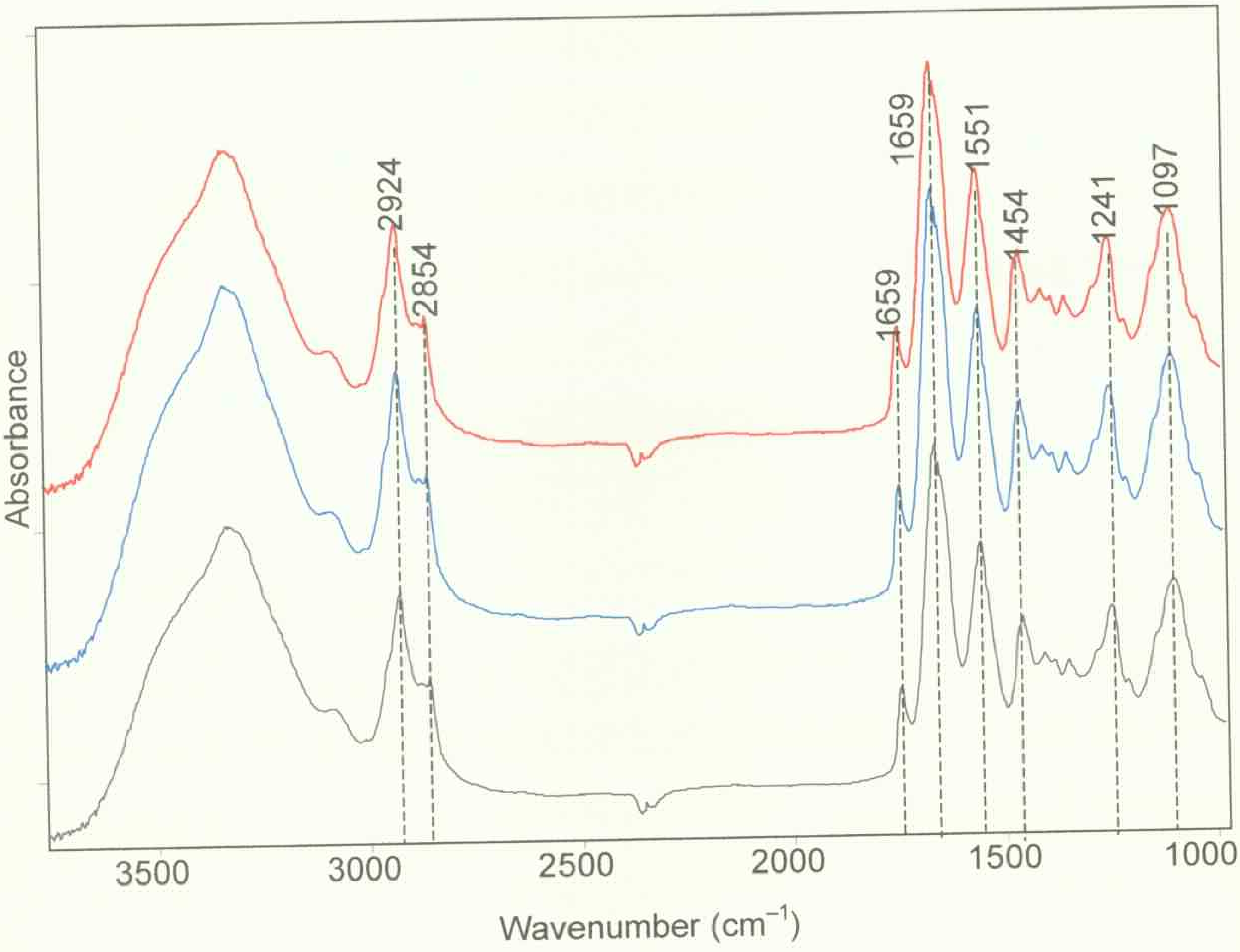


Figure 4.13. Spectral repeatability of an invasive ductal carcinoma breast cryosection spectral collected at different locations from tumour cells within the same cancerous duct. Experimental conditions: 512 scans; 4 cm^{-1}; ×15 magnification; and average of triplicates. Spectra are offset for clarity.

Spectral variations among the formalin-fixed cryosections in the PC scores plot (Figure 4.7) were due to the heterogeneity of the breast tissue. To illustrate the inherent heterogeneity of the sample, Figure 4.14 shows four representative IR spectra measured at points corresponding to various tissue components from an infiltrating lobular carcinoma cryosection. The H&E stained section was used to locate the major breast tissue components for IR spectral collection on an adjacent unstained section. As seen here, spectra collected at the various locations produced different spectra, as would be expected. In addition, most of the breast tissue components (ducts, lobules and adipose tissues) vary in size and the use of inappropriate aperture sizes (5 × 5 mm^2 in a typical IR microscope) could result in data collection of an area containing too much or too little of the component of interest. This can lead to a false positive or negative for diagnostics. However, the S/N becomes poorer as the consequence of closing the aperture to collect data from a specific area, which could be solved/improved by using SR-FTIR microspectroscopy [15,16]. The advantages of SR-FTIR spectroscopy will be discussed in Chapter 5, Section 5.4.

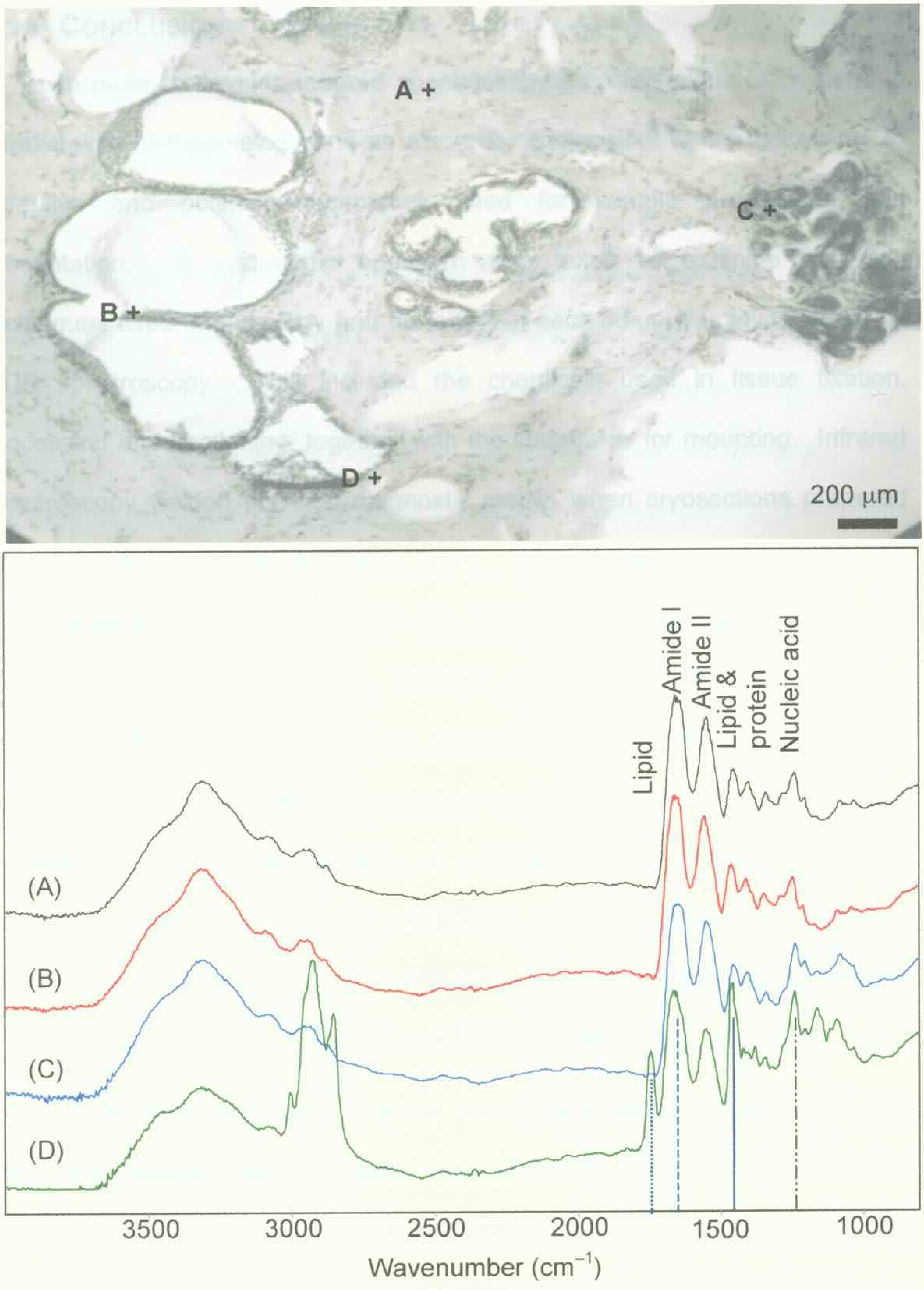


Figure 4.14. (Top) Photomicrograph of a 5-µm-thick H&E stained invasive lobular carcinoma cryosection (adipose removed during staining process). (Bottom) IR spectra normalised to the amide I band of an adjacent 5-µm-thick cryosection measured at various locations: (A) connective tissues (stroma); (B) nuclei at duct; (C) nuclei at lobules; and (D) nuclei surrounded by adipose tissues. Experimental conditions: 512 scans; 4 cm^{-1}; ×15 magnification; transflection mode using transflective slide. Spectra are offset for clarity.

4.11 CONCLUSION

In order to develop infrared spectroscopy as a technique to be used in parallel with histopathology and as a tool for diagnosis it is first necessary to establish and optimise techniques used for sample preparation and presentation. A systematic approach was taken to examine sampling techniques used in pathology and the effect of each stage was investigated by FTIR spectroscopy. This included the chemicals used in tissue fixation, processing and sectioning, together with the substrates for mounting. Infrared spectroscopy yielded the best diagnostic results when cryosections prepared from formalin-fixed tissues were used. This treatment preserved the morphological structure, minimised the effect of autolysis, and allowed the option of retrospective analysis. Principal component analysis was used to separate different stages of tissue processing using the 3050–2750 cm^{-1} spectral region, which is found to be sensitive to the chemicals used in tissue processing. This allows samples of unknown processing to be classified according to their processing and sectioning. The immediately cryosectioned biopsy was clearly separated from the same biopsy that was fixed in formalin and microtomed either as cryosection or FFPE section.

Cover slips are used to avoid tissue degradation on stained tissue sections in histopathology, however, most of the IR studies to date place the tissue sections (cryosections/FFPE sections) on infrared mounting substrates uncovered [13,17-29]. This raises the question of tissue stability. ATR spectra of two malignant biopsies that were fixed in formalin solution for three and half years were compared with the original spectra obtained from cryosections*. No protein degradation was observed in the formalin-fixed biopsies, however,

---

* Insufficient sample volume remained for cryosections to be prepared from the three-and-half year old biopsy.

significant shifts in the positions of the amide I and II bands were noted and were ascribable to protein cross-linking due to formalin fixation as previously assigned in the literature [6,8]. The biopsy stored in PBS/D_2O displayed signs of cellular disintegration and retrospective analysis could not be undertaken.

Although the anomalous dispersion artefact was observed within the 2000–1700 cm^{-1} spectral region, the transflective slide was determined to be the best mounting substrate. The samples on the substrate produced high quality spectra and the sections could be stained after IR analysis. ATR spectroscopy is a relatively simple technique, where no sample preparation was required. This was particularly useful for measuring fresh or fixed tissue biopsies, biological references and chemical solutions. However, the use of ATR for analysis of bulk tissue and sections is not recommended for the following reasons. The spectra of water and formalin can interfere with that of the tissue if analysing bulk tissue biopsies. Furthermore, tissue sections are in physical contact with the ATR crystal surface and could be damaged, there is also the likelihood of contamination with other chemicals and sections as seen in Figures 4.2 and 4.9. IR microspectroscopy was chosen to be the best sampling technique for tissue section analysis because individual tissue components (ducts, lobules, connective and adipose tissues) could be located using the adjacent stained section.

To summarise, the optimal sampling conditions were found to be – a 5 to 10 μm cryosection, mounted on a transflective slide with the proviso that the sample was measured within a week of sample acquisition. In addition, the adipose content remains during the cryosectioning process. However, if the samples are required for retrospective analysis, a 5 μm-thick formalin-fixed paraffin-embedded section with proper deparaffinisation was prepared. This prevents the tissue undergoing autolysis and putrefaction.

4.12 REFERENCES

[1] G. Socrates, *Infrared and Raman Characteristic Group Frequencies: Tables and Charts*. 3rd ed., John Wiley & Sons Ltd., Chichester, 2001.

[2] G. McKee, *Cytopathology of the Breast*. 1st ed., Oxford University Press, Boston, 2002.

[3] J. Radford, *Histopathology Techniques*. 7th ed., The University of Sydney, Sydney, 2002.

[4] P. Rosen, *Rosen's Breast Pathology*. 2nd ed., Lippincott Williams & Wilkins, Philadelphia, 2001.

[5] C. Liu, Y. Zhang, X. Yan, X. Zhang, C. Li, W. Yang, D. Shi, *J. Lumin.* 119-120 (2006) 132-136.

[6] E. Faolain, M. Hunter, J. Byrne, P. Kelehan, M. McNamara, H. Byrne, F. Lyng, *Vib. Spectrosc.* 38 (2005) 121-127.

[7] R. Dukor, *Vibrational Spectroscopy in the Detection of Cancer in Handbook of Vibrational Spectroscopy*. 1st ed., John Wiley & Sons, Chichester, 2002, Vol. 5, 3335-3361.

[8] N. Pleshko, A. LBoskey, R. Mendelsohn, *Calcif. Tissue Int.* 51 (1992) 72-77.

[9] R. Blamey, A. Evans, I. Ellis, R. Wilson, *Atlas of Breast Cancer*. 1st ed., Merit Publishing International, Basingstoke, 1994.

[10] S. Silverberg, *Atlas of Breast Pathology*. Elsevier Science (USA), Philadelphia, 2002.

[11] J. Underwood, *General and Systematic Pathology*. 2nd ed., Churchill Livingstone, New York, 1996.

[12] N. Weidner, *Breast in Modern Surgical Pathology*. Saunders, New York, 2003, Vol. 1, 540-571.

[13] M. Diem, M. Romeo, *Vib. Spectrosc.* 38 (2005) 115-119.

[14] M. Romeo, M. Diem, *Vib. Spectrosc.* 38 (2005) 129-132.

[15] L. Miller, P. Dumas, *Biochem. Biophys. Acta* 1758 (2006) 846-857.

[16] L. Miller, *J. Biol. Phys.* 29 (2003) 219-230.

[17] G. Yu, J. Xu, Y. Niu, C. Zhang, C. Zhang, *Proc. SPIE-Int. Soc. Opt. Eng.* 5630 (2005) 796-801.

[18] G. Yu, S. Lu, J. Xu, C. Zhang, G. Zhang, *Nankai Daxue Xuebao, Ziran Kexueban* 38 (2005) 53-56.

[19] G. Yu, S. Lu, J. Xu, C. Zhang, C. Zhang, *Guangzi Xuebao* 34 (2005) 390-394.

[20] Y. Xu, Y. Zhao, Z. Xu, Y. Ren, Y. Liu, Y. Zhang, X. Zhou, J. Shi, D. Xu, J. Wu, *Guangpuxue Yu Guangpu Fenxi* 25 (2005) 1775-1778.

[21] D. Wetzel, G. Post, R. Lodder, *Vib. Spectrosc.* 38 (2005) 53-59.

[22] E. Gazi, J. Dwyer, N. P. Lockyer, J. Miyan, P. Gardner, C. A. Hart, M. D. Brown, N. W. Clarke, *Vib. Spectrosc.* 38 (2005) 193-201.

[23] E. Gazi, J. Dwyer, N. P. Lockyer, J. Miyan, P. Gardner, C. Hart, M. Brown, N. W. Clarke, *Biopolymers* 77 (2005) 18-30.

[24] E. O. Faolain, M. B. Hunter, J. M. Byrne, P. Kelehan, H. J. Byrne, F. M. Lyng, *Proc. SPIE-Int. Soc. Opt. Eng.* 5826 (2005) 25-36.

[25] B. Wood, L. Chiriboga, H. Yee, M. A. Quinn, D. McNaughton, M. Diem, *Gynecol. Oncol.* 93 (2004) 59-68.

[26] Q.-B. Li, Z. Xu, Y.-Z. Xu, Y.-F. Zhang, N.-W. Zhang, L.-X. Wang, X.-J. Sun, L. Zhang, F. Wang, L.-M. Yang, Y. Zhao, Y. Ren, L. Zhi, S.-F. Weng, W.-J. Zhou, J.-G. Wu, *Gaodeng Xuexiao Huaxue Xuebao* 25 (2004) 2010-2012.

[27] P. Lasch, W. Haensch, D. Naumann, M. Diem, *Biochem. Biophys. Acta* 1688 (2004) 176-186.

[28] M. Diem, M. Romeo, C. Matthaus, M. Miljkovic, L. Miller, P. Lasch, *Infrared Phys. Techn.* 45 (2004) 331-338.

[29] K. Bambery, B. Wood, M. A. Quinn, D. McNaughton, *Aust. J. Chem.* 57 (2004) 1139-1143.

5. CHARACTERISATION AND DIFFERENTIATION OF BREAST TISSUES USING INFRARED SPECTROSCOPY

5.1 SPECTRAL INTERPRETATION OF BREAST TISSUE

Breast tissue is heterogeneous and contains various tissue components such as lobules, ducts, and adipose and connective tissues. Each tissue type contains characteristic concentrations of cellular and extracellular components composed of proteins (including extracellular structural proteins, such as collagens), nucleic acids, lipids, etc. [1]. The mid-infrared (MIR) spectral region (4000–650 $\mathrm{cm^{-1}}$) provides a wealth of information about the relative amounts of these biological building blocks present in a sample [2-28].

Infrared spectra of biological samples contain numerous overlapping bands due to the presence of many biochemicals with similar functionality [2-4]. Table 5.1 lists the commercially available biochemical references (Sigma-Aldrich Co. Ltd., USA) that were used to help delineate the main features of tissue spectra.

Table 5.1. Biochemical references used for identification of breast tissue components.

Chemicals	Form	Biological References
Calf thymus DNA	Solid	DNA
Mucin Type I-S from bovine submaxillary glands	Solid	Protein
Bovine albumin	Solid	Protein
Oleic acid	Solution	Fatty acid

Figure 5.1 presents the infrared spectra of tumour cells found around the duct and surrounded by adipose tissues of an invasive ductal carcinoma (IDC) breast cryosection together with the biochemical reference samples used in this research. The band assignments for the spectrum of the IDC cryosection will be discussed in three regions: (I) 4000–3050 $\mathrm{cm^{-1}}$; (II) 3050–2750 $\mathrm{cm^{-1}}$; and (III) 1800–650 $\mathrm{cm^{-1}}$. The 2750–1800 $\mathrm{cm^{-1}}$ region contains no significant diagnostic information, i.e., biomolecules do not usually absorb in this region.

The major bands and their assignments are summarised in Table 5.2.

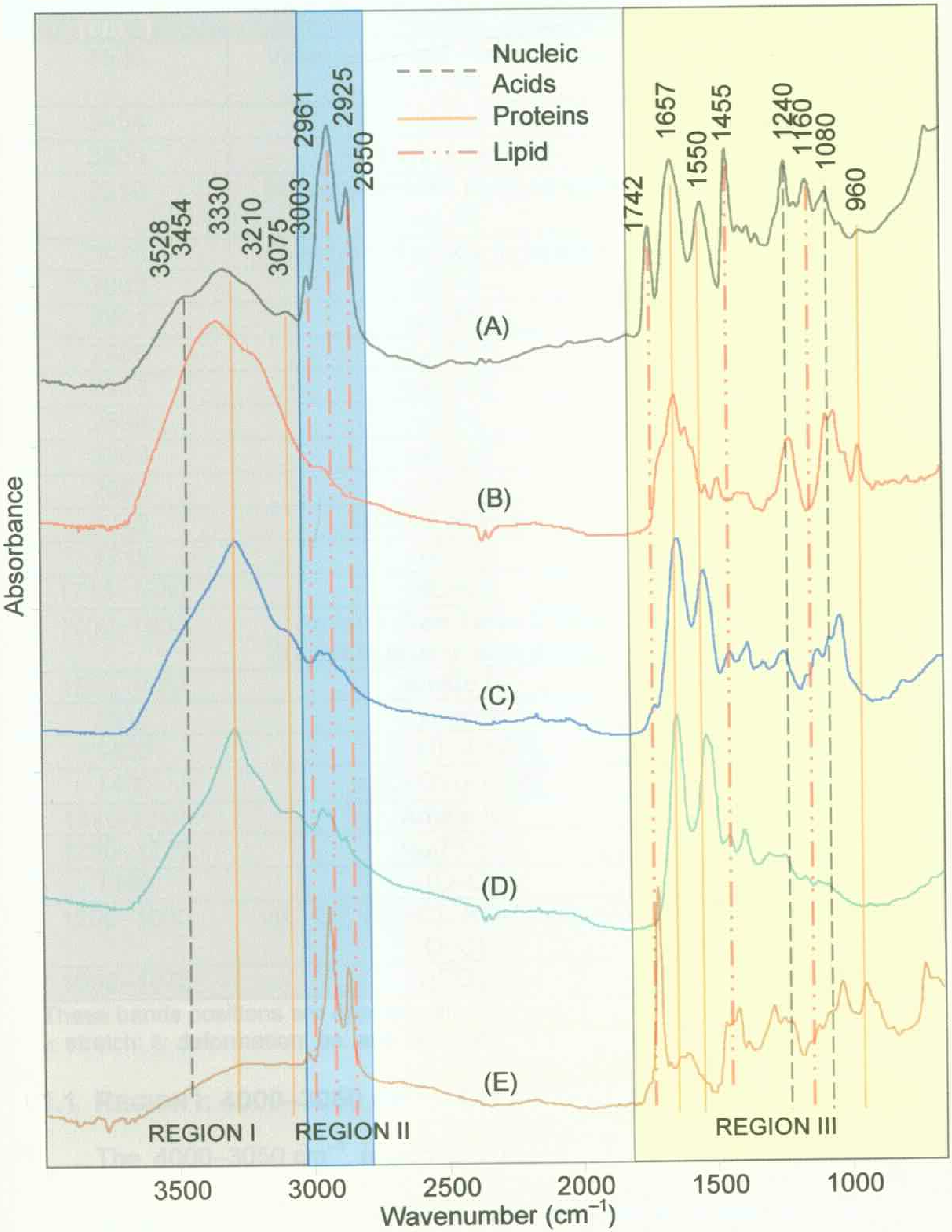


Figure 5.1. Comparison of infrared spectra of: (A) a 5-μm-thick malignant IDC breast cryosection (area containing tumour cells and adipose tissues) and commercially available biochemical references: (B) DNA; (C) mucin type I-s glycoprotein; (D) bovine albumin protein; and (E) oleic acid. Experimental conditions: (A) transflection spectrum (512 scans; 4 cm^{-1}; x15 magnification; and average of triplicates); and (B–E) ATR spectra (512 scans; 4 cm^{-1}; and average of triplicates). Specific lines highlight the bands from major biological components. Spectra are to scale and offset for clarity.

Table 5.2. IR band assignments of breast tissue and biological references [1,4-7].

Wavenumber* ($\mathrm{cm^{-1}}$)	Assignments#	Biological Reference
3528	Weakly-bonded residual water $\nu(\mathrm{OH})$	Water
3454	$\nu(\mathrm{NH_2})$	Nucleic acid
3330	$\nu(\mathrm{N{-}H})$ (amide A)	Proteins
3210	Strongly-bonded residual water $\nu(\mathrm{OH})$	Water
3075	Overtone of amide II (amide B)	Proteins
3003	$\nu(\mathrm{CH})$	Lipid
2961	$\nu_s(\mathrm{CH_3})$	Aliphatic
2925	$\nu_{as}(\mathrm{CH_2})$	Aliphatic
2918	$\nu_{as}(\mathrm{CH_2})$	Fatty acids
2898	$\nu(\mathrm{CH})$	Methine groups
2868	$\nu_s(\mathrm{CH_3})$	Aliphatic
2850	$\nu_s(\mathrm{CH_2})$	Fatty acids
1742	$\nu(\mathrm{C{=}O})$ of esters	Fatty acids
1715	$\nu(\mathrm{C{=}O})$	Carbohydrates
1715–1680	$\nu(\mathrm{C{=}O})$	Nucleic acids
1700–1635	Amide I (See Table 5.3 for complete protein assignments)	Proteins
1550-1520	Amide II	Amino acids
1515	Tyrosine	Amino acids
1455	$\delta(\mathrm{C{-}H})$ of $>\mathrm{CH_2}$	Mainly fatty acids
1400	$\nu_s(\mathrm{C{=}O})$ of $\mathrm{COO^-}$	Carbohydrates
1310-1240	Amide III	Proteins
1250–1220	$\nu_{as}(\mathrm{PO_2^-})$	Nucleic acids
1160	$\nu(\mathrm{C{-}C})$	Lipid, carbohydrates
1200–1000	$\nu(\mathrm{C{-}O})$, $\nu(\mathrm{C{-}C})$, $\delta(\mathrm{C{-}O{-}H})$, $\delta(\mathrm{C{-}O{-}C})$	Carbohydrates
1090–1075	$\nu_s(\mathrm{PO_2^-})$	Nucleic acids

* These bands positions are approximate only and may shift ±10 $\mathrm{cm^{-1}}$ or more.

\# ν: stretch; δ: deformation; as: anti-symmetric; s: symmetric.

5.1.1 REGION I: 4000–3050 CM^{-1} – OH AND NH STRETCHING VIBRATIONS

The 4000–3050 $\mathrm{cm^{-1}}$ region is comprised of a number of overlapping bands assigned to the $\nu(\mathrm{OH})$ mode of water and carbohydrates and the $\nu(\mathrm{NH_2})$ and $\nu(\mathrm{NH})$ modes of nucleic acids and proteins, respectively [4-6,8]. The spectral region of the IDC breast cryosection (Figure 5.1A) has been curve fitted, using the second-derivative to locate the positions of the bands within the spectral envelope, these results are presented in Figure 5.2.

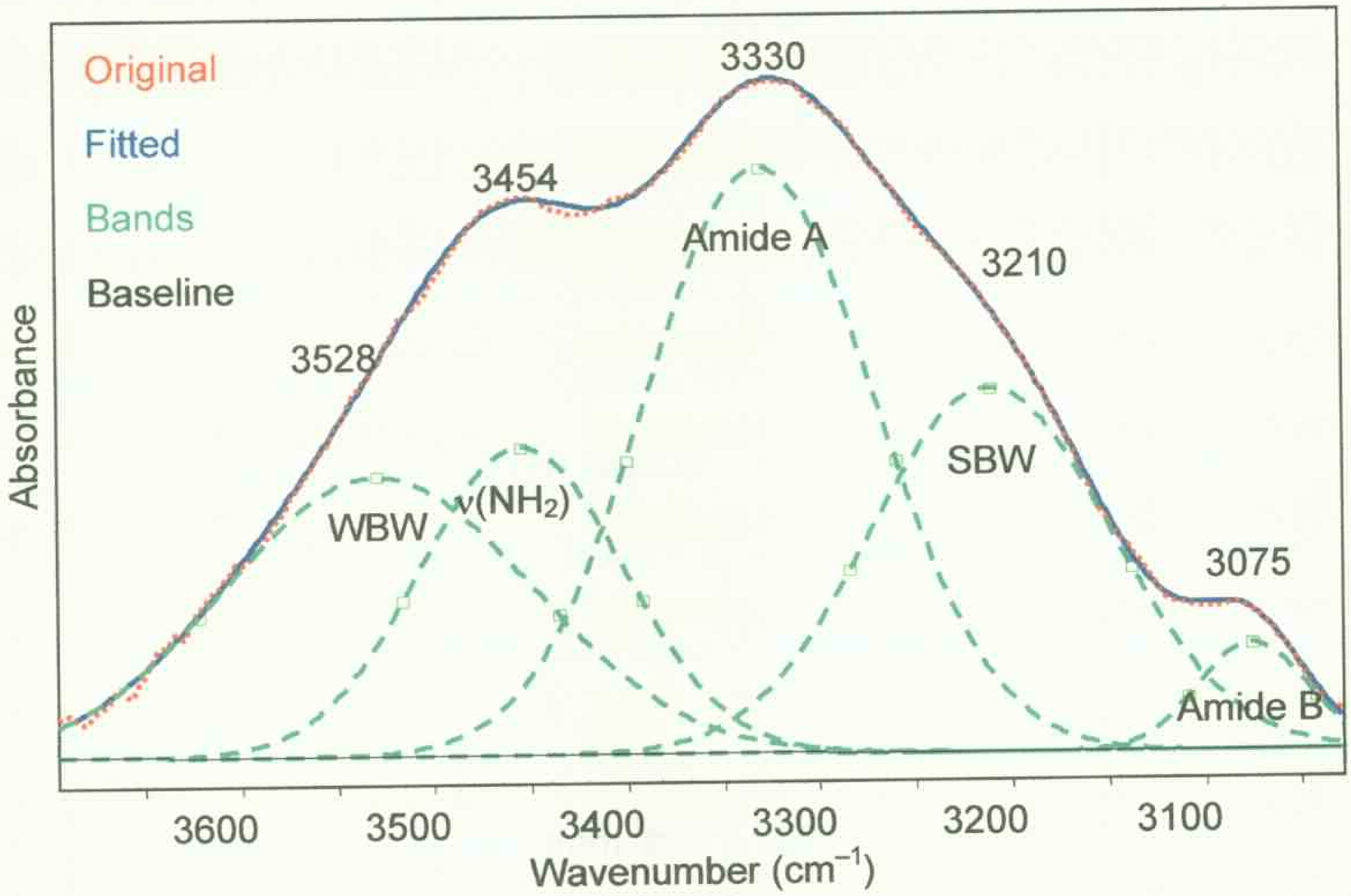


Figure 5.2. Region I: 4000–3050 cm^{-1} region from the FTIR spectrum of a malignant IDC breast cryosection (Figure 5.1A) showing the original trace (red); curve-fitted trace (blue); and the individual resolved bands within this region (green). Prior to iteration, the positions of the individual bands were located using the Savitzky-Golay second-derivative method.

The broad ν(OH) absorptions centred at approximately 3528 and 3210 cm^{-1} are due to weakly (WBW) and strongly bonded (SBW) residual water, respectively, involved in hydrogen bonding within the tissue [5,6,9]. In the majority of published research of this nature, standard PBS/H_2O has been used to maintain the physiological pH of tissues prior to sectioning [17,18,20-22,26,29,30]. However, breast biopsies used in this research were stored in PBS/D_2O for parallel studies with MRS, which were washed with distilled H_2O during tissue processing (both cryosectioning and FFPE sectioning). The spectral effect of the H_2O/D_2O exchange that occurred during tissue processing was discussed in detail in Chapter 4, section 4.2.

To briefly summarise, D_2O was observed to have an intense IR absorption in the cryosection spectrum (Figure 4.2C), which was centred at 2482 cm^{-1}, assigned to the ν(O–D) mode, but this band was absent in the FFPE section spectrum (Figure 4.2B). A small shift in the position of the water ν(O–H) band to a higher wavenumber, from 3324 to 3394 cm^{-1}, was observed in the cryosection spectrum compared to the deparaffinised FFPE spectrum. Therefore, with different sectioning used in the analysis, a small wavenumber shift in the position of the ν(O–H) band, due to D_2O/H_2O exchange, would be expected. Consequently, this spectral region was not routinely used in multivariate statistical analysis for tissue classification of single-point and imaging spectra.

The strong contribution from the protein ν(NH) bands at 3330 and 3075 cm^{-1}, assigned to the amide A and B bands, are evident in the spectrum of the breast tissue sample when it is compared to the spectra obtained from DNA, glycoprotein and bovine albumin protein (Figures 5.1A–D) [31]. The shoulder centred at 3454 cm^{-1} is due to $\nu(NH_2)$ vibrations of nucleic acids [14,31]. The amide B (overtone of amide II) band at 3075 cm^{-1} is absent in the DNA spectrum (Figure 5.1B), since there is no peptide linkage.

5.1.2 REGION II: 3050–2750 CM^{-1} – CH STRETCHING VIBRATIONS

The 3050–2750 cm^{-1} spectral region (Figure 5.3) contains bands that are assigned to various CH symmetric and anti-symmetric stretching vibrations [31]. These bands originate from proteins, lipids and nucleic acids, however, the major contributor to the intensity of these bands are the lipids (Figure 5.1E). Other bands that are also present in this region are attributable to other C–H containing biomolecules and chemicals used in tissue processing (see Section 4.3) [31].

Based on the literature, the bands at 2961 and 2868 cm^{-1} are assigned to $\nu_{as}(CH_3)$ and $\nu_s(CH_3)$ modes, whereas those at 2925 and 2850 cm^{-1} are assigned to $\nu_{as}(CH_2)$ and $\nu_s(CH_2)$ modes, respectively [4,8,10]. The band at 3003 cm^{-1} is ascribed to the ν(CH) mode of aromatic rings [31].

The relative intensity and lineshape of this spectral region is highly dependent upon the amount of the individual tissue components present in the sample. The very intense $\nu_{as}(CH_2)$ and $\nu_s(CH_2)$ bands at 2923 and 2850 cm^{-1}, respectively, in the IDC spectrum (Figure 5.1A) are indicative of a high lipid content; the major sources of lipids being adipose tissues and biological membranes [32,33]. It is proposed that the intensities of these bands can be used to locate the intra-cellular and extra-cellular components present in different regions of the breast tissue.

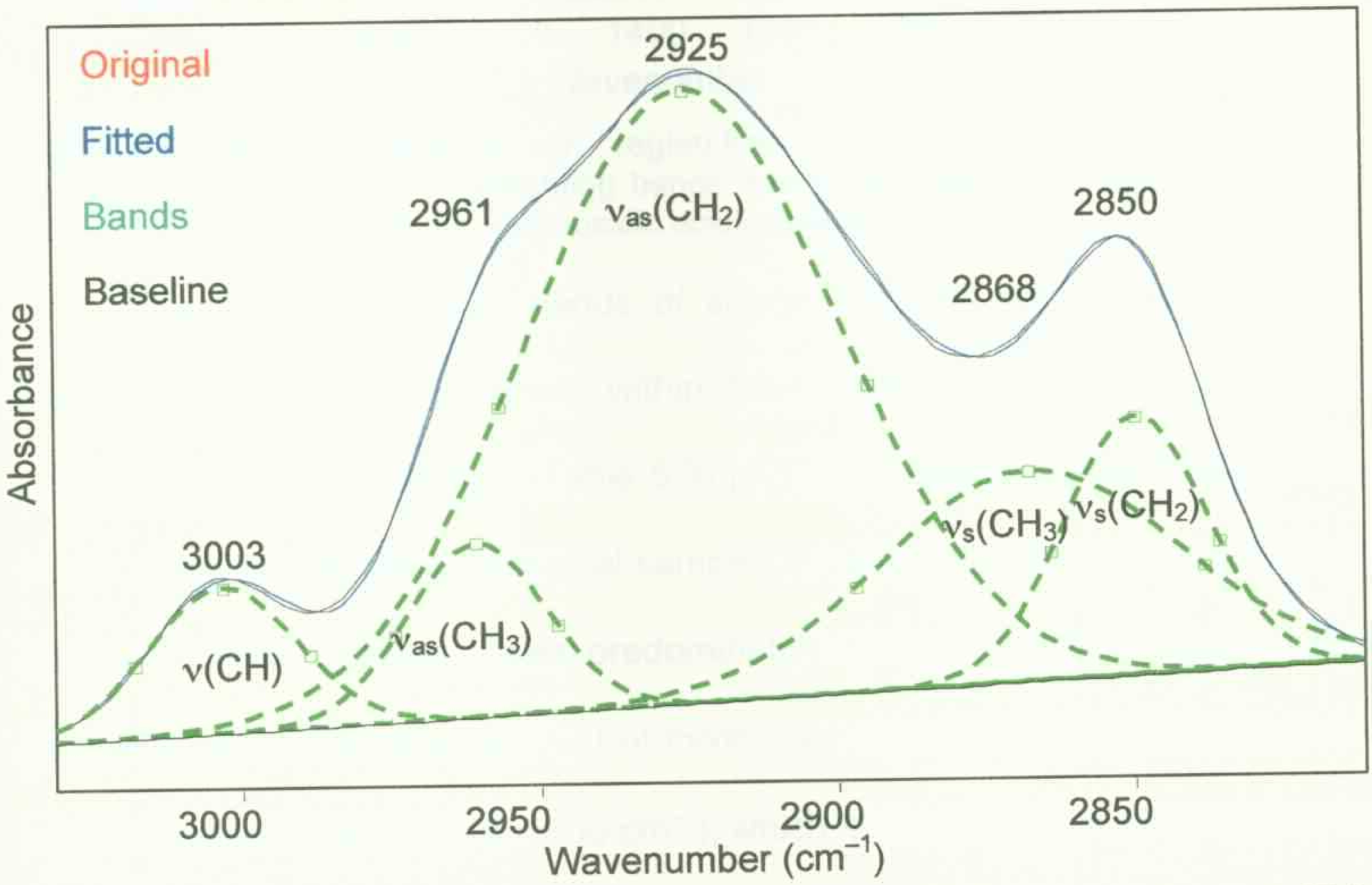


Figure 5.3. Region II: 3050–2750 cm^{-1} region from the FTIR spectrum from an IDC breast cryosection in Figure 5.1A showing spectra of original trace (red); fitted trace (blue); and individual resolved bands under the fitted trace (green). The positions of individual bands were determined using the Savitzky-Golay second-derivative method prior to iterative fitting of the overlapping bands in this region.

5.1.3 REGION III: 1800–650 CM^{-1} – FINGERPRINT REGION

The 1800–650 cm^{-1} region (Figure 5.4) is referred to as the 'fingerprint' region, and contains diagnostic infrared absorptions from proteins, lipids, nucleic acids, carbohydrates and other biomolecules [8,31].

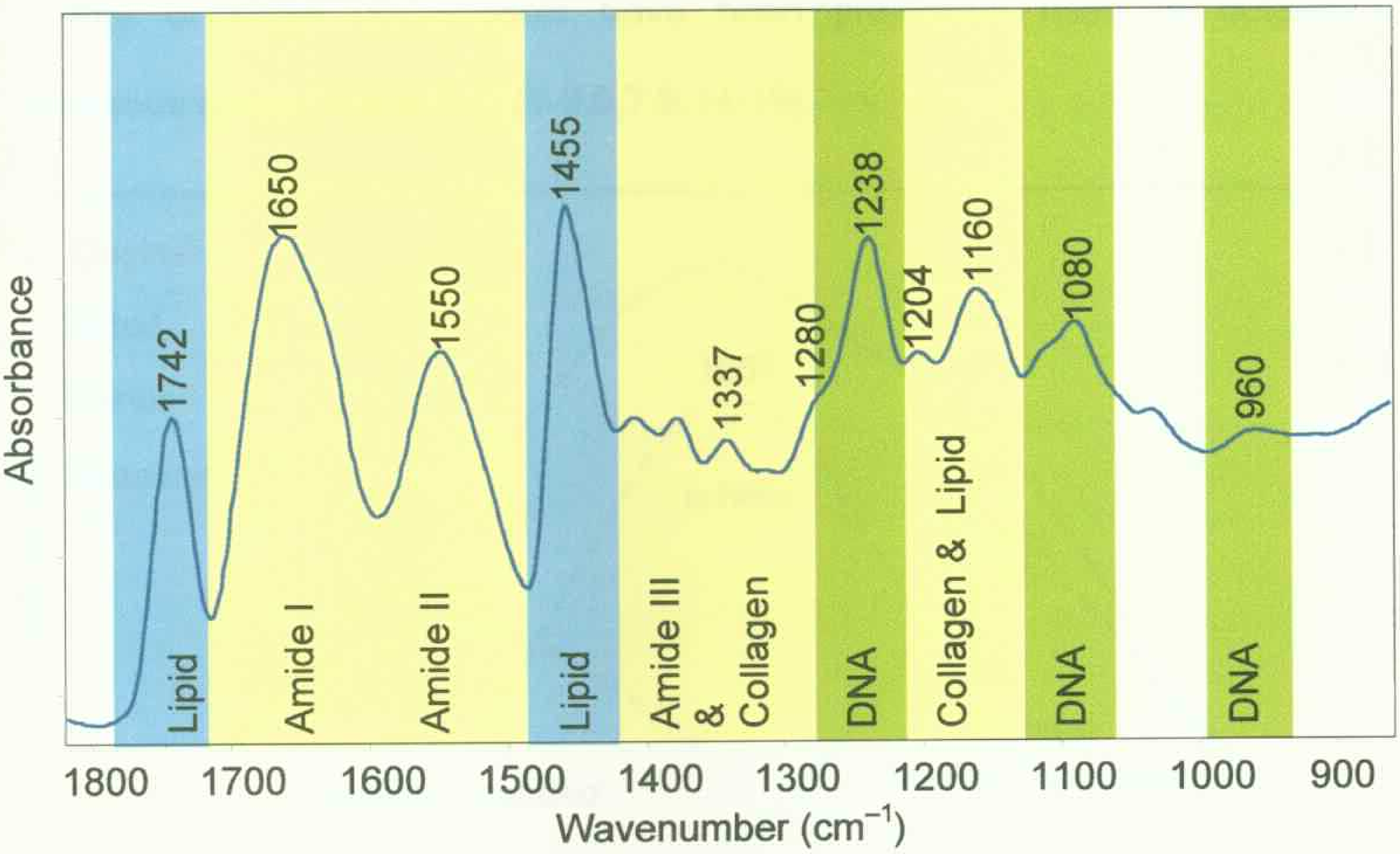


Figure 5.4. Region III: 1800–650 cm^{-1} region from the FTIR spectrum of an IDC breast cryosection in Figure 5.1A highlighting bands due to biomolecules in specific colours: lipids (blue); proteins (yellow); and nucleic acids (green) [31].

There are two amide bands of strong intensity, and numerous other weaker amide bands observed within the fingerprint region, all of which originate from peptide bonds (Table 5.3) [12]. The stronger absorption bands found in the IR spectra of biological samples (Figure 5.4) are: the amide I band (1695–1610 cm^{-1}), which arises predominately from the ν(C=O) mode of the peptide bond with a small amount of mixing with the ν(CN) and δ(NH) modes; and the amide II band (1575–1480 cm^{-1}), which is assigned to a combination of ν(CN) and δ(NH) modes [12,31]. The amide III band (1220–1320 cm^{-1}) arises from a combination of ν(CN), δ(NH), ν(C=O) and δ(O=C–N) modes and is relatively weaker in intensity compared with the amide I and II bands [12,31].

The position of the amide I band is sensitive to the proportions of different protein secondary structures (e.g., α-helix, β-sheet and random coil) present in cells/tissues [8,12,29]. Curve fitting of the amide I region reveals numerous individual bands that contribute to this region (Figure 5.5). The position and amount of these components have been previously used to determine cancerous states (Table 5.3) [1-3,5,7,9,14-19].

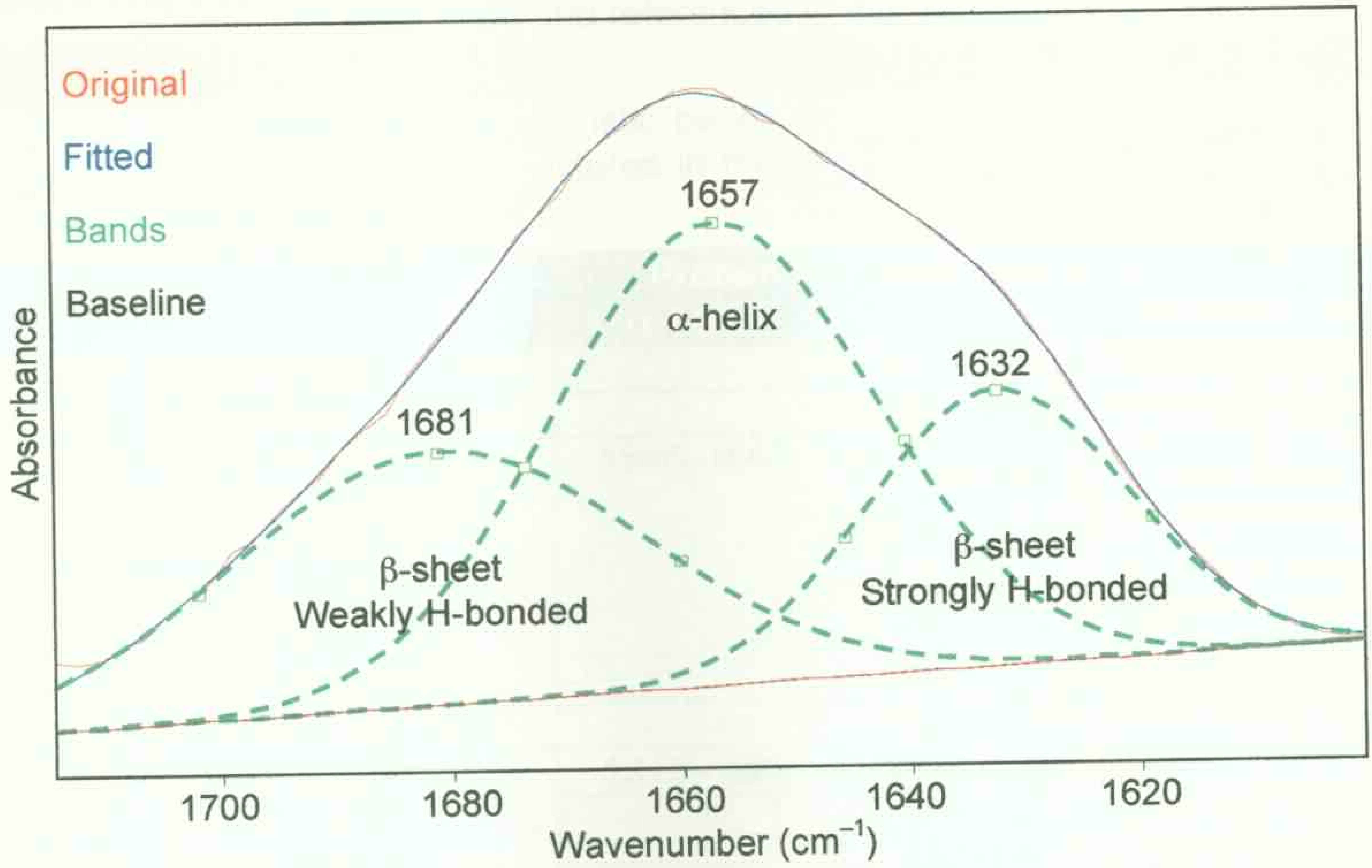


Figure 5.5. Curve-fitted amide I band from the FTIR spectrum of the IDC breast cryosection in Figure 5.1A showing the major bands due to the various secondary structures including weakly- and strongly-H-bonded β-sheet components and the α-helix band. The positions of individual bands were determined using the Savitzky-Golay second-derivative method prior to iterative fitting of the overlapping bands.

Collagen is the major protein component found in connective (stromal) tissues; it contains approximately 33% glycine and 20–25% proline and hydroxyproline residues [8,31]. Infrared bands assigned to collagen include the amide I band at ~1650 cm^{-1}, amide II band at 1550 cm^{-1} and amide III bands at 1337, 1280, 1237 and 1204 cm^{-1} [8,31]. The δ(CH_2) and δ(CH_3) bands occur at 1450 and 1401 cm^{-1} and the collagen carbohydrate residues result in bands at 1080 and 1204 cm^{-1} [8,31]. The bands assigned to the amide III vibrations also have contributions from the wagging vibrations of glycine backbones and proline side-chains [8]. The amide I functional group maps plotted from FTIR imaging and SR-FTIR mapping experiments, which will be discussed in detail later in the chapter, therefore, comprise of both proteins from epithelial cells and collagens from stromal tissues. In order to differentiate these two components, multivariate statistical analysis, including k-means clustering and fuzzy c-means clustering analyses were used and referenced to the histopathology results.

Table 5.3. Positions of characteristic bands for amide vibrational modes and their associated secondary structures in the MIR region [12,31]. w: weak; m: medium; and s: strong.

Amide	Intensity	Protein Secondary Structure (cm^{-1})		
		α-helix	β-sheet	Random Coil
A	m-s	~3300		
B	w	~3100		
I	s	1660–1645	1680–1665 (weakly H-bonded) 1640–1620 (strongly H-bonded)	1670–1660 (overlapped with weakly H-bonded β-sheet)
II	s	~1550		
III	m-w	1310–1260	1240–1225	1260–1240
IV	w	765–625		
V	w	800–640		
VI	w	605–535		
VII	w	~200		

The prominent nucleic acids bands (mainly DNA) are observed at 1240 and 1080 cm^{-1} and are attributed to the anti-symmetric and symmetric stretching modes of the phosphodiester backbones, $\nu_{as}(PO_2^-)$ and $\nu_s(PO_2^-)$, respectively [4,16,20]. DNA also has bands due to C=O and aromatic C=C stretching vibrations at 1688, 1643, 1603 and 1577 cm^{-1} as shown in Figure 5.1B [4,16,20].

Infrared absorption bands of lipids found in this region (Figure 5.4) include those due to: the lipid-ester $\nu(C{=}O)$ mode at 1742 cm^{-1}; the $\delta(CH_2)$ modes of long-chain hydrocarbons at 1455 cm^{-1}; and the $\nu(C{-}C)$ mode at 1160 cm^{-1} [21-25]. The intensity of the $\nu(C{=}O)$ lipid ester band varies depending on the adipose levels in the tissue section. In this work, the spectra of breast cryosections (FFPE sectioning removes adipose content during tissue processing) with high adipose content were found to have a very intense lipid ester $\nu(C{=}O)$ band together with very intense $\nu_{as}(CH_2)$ and $\nu_s(CH_2)$ bands in Region II, and is consistent with the findings from literature [8,18,29,34-36].

5.2 DIFFERENTIATION OF BREAST LESIONS USING SINGLE-POINT MICROSPECTROSCOPY

Lobules, ducts, adipose and stromal tissues are the major components of breast tissue [32,33]. The lobules and ducts are comprised of outer myoepithelial and inner cuboidal epithelial cells that are surrounded by basement membranes, and are supported by the adipose and stromal tissues [32,33]. Fibrocystic changes are the most common form of breast lesion and arise due to the formation of cysts within the duct [32,33]. No cellular proliferation is found in this condition as described in Section 3.3.1. Tumour cells within the ducts or lobules are associated with the development of breast cancer. These proliferations result in an increase of abnormal nuclei/cells, which may invade the basement membranes of the ductal and lobular epithelial cells. If not diagnosed early, the tumour cells may metastasise to other part(s) of the body [11,26-28]. In the research described in Chapter Three, it has been shown that different breast diseases displayed various morphological features that could be distinguished in the H&E stained sections using light microscopy. These morphological transformations result in elevated intensities of nucleic acid and protein bands and are attributed to the tumour cells [1,7,14,18,29-31]. The lipid levels are also expected to increase due to the additional cell membrane material [1,7,14,18,29-31]. Therefore, by comparing the intensities of these characteristic IR bands, it is possible to differentiate and characterise breast lesions.

Representative infrared spectra from: (A) a normal breast duct; three benign breast lesions: (B) fibrocystic changes; (C) fibroadenoma; (D) intraduct papilloma; two *in-situ* carcinomas: (E) ductal carcinoma *in-situ* (DCIS); (F) lobular carcinoma *in-situ* (LCIS); and two malignant lesions: (G) invasive ductal carcinoma (IDC); and (H) invasive lobular carcinoma (ILC), are shown in Figure 5.6. All spectra were acquired from the ducts or lobules, which are the major tissue component(s) of each lesion. The FTIR microscope used to obtain the spectra in this research has a knife-edge aperture with dimensions of 5.0×5.0 mm^2, however, the size of the major tissue components ranged from 100 μm (lobules) to over 1 mm (ducts) in diameter. Although the knife-edge aperture could be reduced to mask only the specific areas of interest, the inherent complexity of the sample led to single-point spectra containing signals originating from more than one tissue component. Furthermore, a reduction in the aperture size resulted in spectra with a significant reduction in the S/N ratio as a consequence of reduced signal throughput [23,37].

It has been demonstrated that spectra from samples prepared using different tissue presentations, i.e., cryosection or formalin-fixed paraffin embedded (FFPE) section, had additional spectral features due to the chemicals used for processing (Chapter 4, Sections 4.3 and 4.4). Therefore, spectral comparisons were based on cryosections of the various lesions. The only exception was the DCIS biopsy, which was acquired embedded in a paraffin block and was subsequently prepared as a deparaffinised FFPE section.

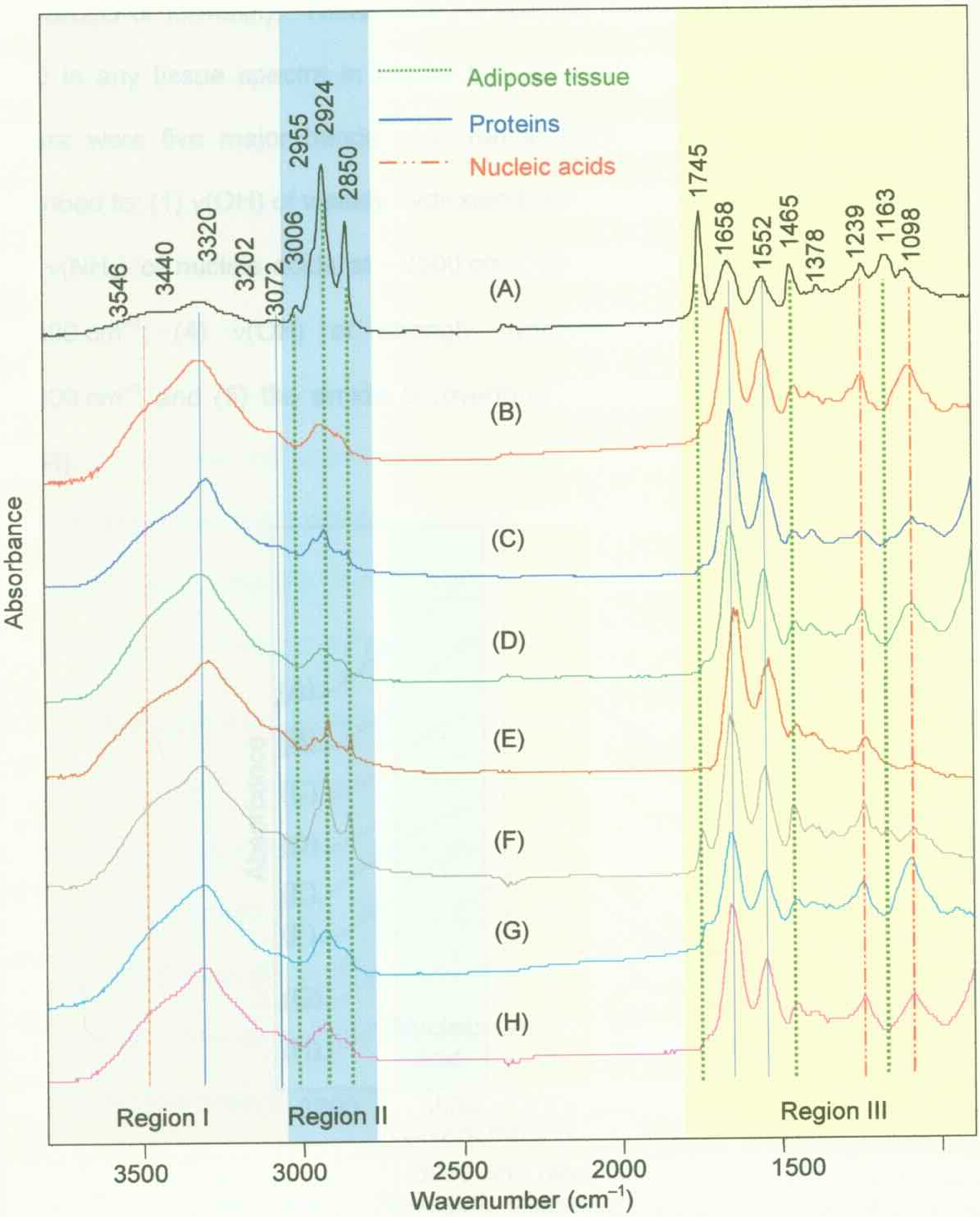


Figure 5.6. Representative infrared spectra of various breast conditions: (A) normal breast duct; (B) fibrocystic change; (C) fibroadenoma; (D) intraduct papilloma; (E) DCIS; (F) LCIS; (G) IDC; and (H) ILC. Experimental conditions: 512 scans; ×15 magnification, 4 cm^{-1}; and average of triplicates. All sections presented here were cryosections except DCIS, which was a deparaffinised FFPE section. Spectra are to scale and offset for clarity.

5.2.1 REGION I: 3600–3050 CM^{-1}

The samples were air-dried prior to spectral collection in order to minimise the strong ν(OH/OD) absorptions from the storage solutions (PBS/D_2O or formalin). There were no spectral indications of the presence of D_2O in any tissue spectra in Figure 5.6, as was discussed in Section 5.1.1. There were five major bands observed in this spectral region (Figure 5.7) ascribed to: (1) ν(OH) of weakly hydrogen-bonded residual water at ~3550 cm^{-1}; (2) ν(NH_2) of nucleic acids at ~3500 cm^{-1}; (3) ν(NH) of the amide A band at ~3300 cm^{-1}; (4) ν(OH) of strongly hydrogen-bonded residual water at ~3200 cm^{-1} and (5) the amide B (overtone of amide II) band at ~3080 cm^{-1} [6,34].

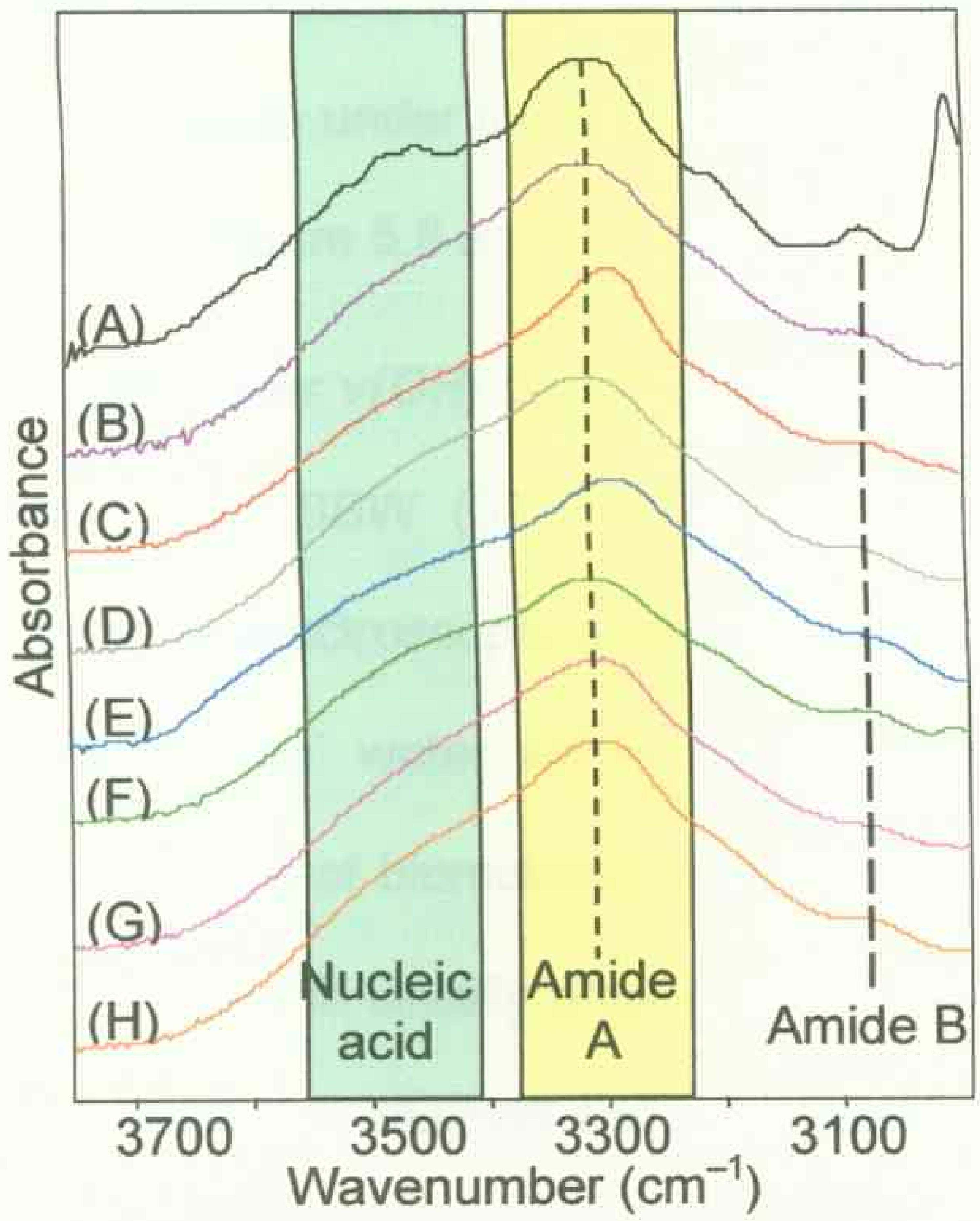


Figure 5.7. Expansion of the 3800–3050 cm^{-1} region from the FTIR spectra of Figure 5.6: (A) normal breast duct; (B) fibrocystic changes; (C) fibroadenoma; (D) intraduct papilloma; (E) DCIS; (F) LCIS; (G) IDC; and (H) ILC. Bands are assigned to weakly and strongly bonded residual water within the tissues, ν(NH_2) bands from nucleic acids and the amide A and amide B (overtone of amide II) bands of proteins and nucleic acids.

As shown in the expanded 3700–3000 cm^{-1} spectral region (Figure 5.7), there are substantial differences in the shapes and positions of the ν(NH_2) nucleic acid and amide A bands at approximately 3400 and 3300 cm^{-1}, respectively, for different tissue conditions [9,14]. The water bands observed in this region are due to weak and strong hydrogen-bonded residual water interacting with biomolecules. The intensity of the nucleic acid ν(NH_2) mode is strongest in the spectra of the normal tissue (Figure 5.7A) compared with the spectra of the other conditions. Similar features were observed in FTIR spectra from samples with fibrocystic changes (Figure 5.7B) and intraduct papillomas (Figure 5.7D), while the positions of the amide A band in the fibroadenoma (Figure 5.7C) and DCIS (Figure 5.7E) spectra were clearly distinguished from those of other lesions in terms of shifts to lower wavenumbers. In order to further investigate these differences, the spectra were subjected to curve-fitting techniques to resolve the bands under the broad absorptions in this region and the results are presented in Figure 5.8 and in Table 5.4.

Hydrogen-bonded water ν(OH) bands: Two bands are assigned to the WBW (3611-3471 cm^{-1}) and SBW (3212–3176 cm^{-1}) in the biomolecules. Tissues were dried prior to spectroscopic analyses; therefore, there would be minimal contribution from 'free' water within the thin sections. The WBW usually attaches to the surface of biomolecules as bridges between molecules or subunits, whereas the SBW is usually present within biomolecules bound to bio-macromolecules, nucleotides, charged and polar groups [14].

The deconvoluted WBW band was present at 3607 ± 4 cm^{-1} in the spectrum of the normal tissue, whereas it shifted to lower wavenumber in the spectra of the benign lesions (fibrocystic changes: 3519 ± 8 cm^{-1}; fibroadenoma: 3528 ± 6 cm^{-1}; and intraduct papilloma 3508 ± 2 cm^{-1}) and in the spectra of the *in-situ* carcinomas (DCIS: 3574 ± 4 cm^{-1} and LCIS: 3540 ± 6 cm^{-1}). The WBW ν(OH) bands in the spectrum of an IDC (Figure 5.8G) and an ILC (Figure 5.8H) also shift to lower wavenumbers, at 3475 ± 4 cm^{-1} and 3515 ± 7 cm^{-1}. The position of the fibroadenoma WBW band was consistent with previous research (3519 ± 5 cm^{-1}), while there were approximately 80 and 21 cm^{-1} wavenumber shifts observed in the normal and carcinoma tissues, respectively, compared with the previous work by Gao [14]. The spectra obtained from lesions in this study were collected predominantly from the epithelial cells of the ducts and the lobules, however, Gao did not report the exact locations where the IR spectra were collected in the tissue lesions [14]. Therefore, the difference in the position of the WBW band in spectra collected from normal and carcinoma sections for this research and Gao's research cannot be comprehensively compared.

The SBW band in the spectra obtained from normal breast ducts was observed at 3192 ± 2 cm^{-1}, and the intensity of this band was relatively stronger in all the spectra of the benign sections. The SBW bands in the spectra obtained from the DCIS and LCIS biopsies were found at 3176 ± 5 cm^{-1} and 3209 ± 1 cm^{-1} and these bands in the spectra from the IDCs, and ILCs biopsies were located at 3196 ± 3 cm^{-1} and 3195 ± 3 cm^{-1}, respectively. These findings were within the ranges published by Gao [14].

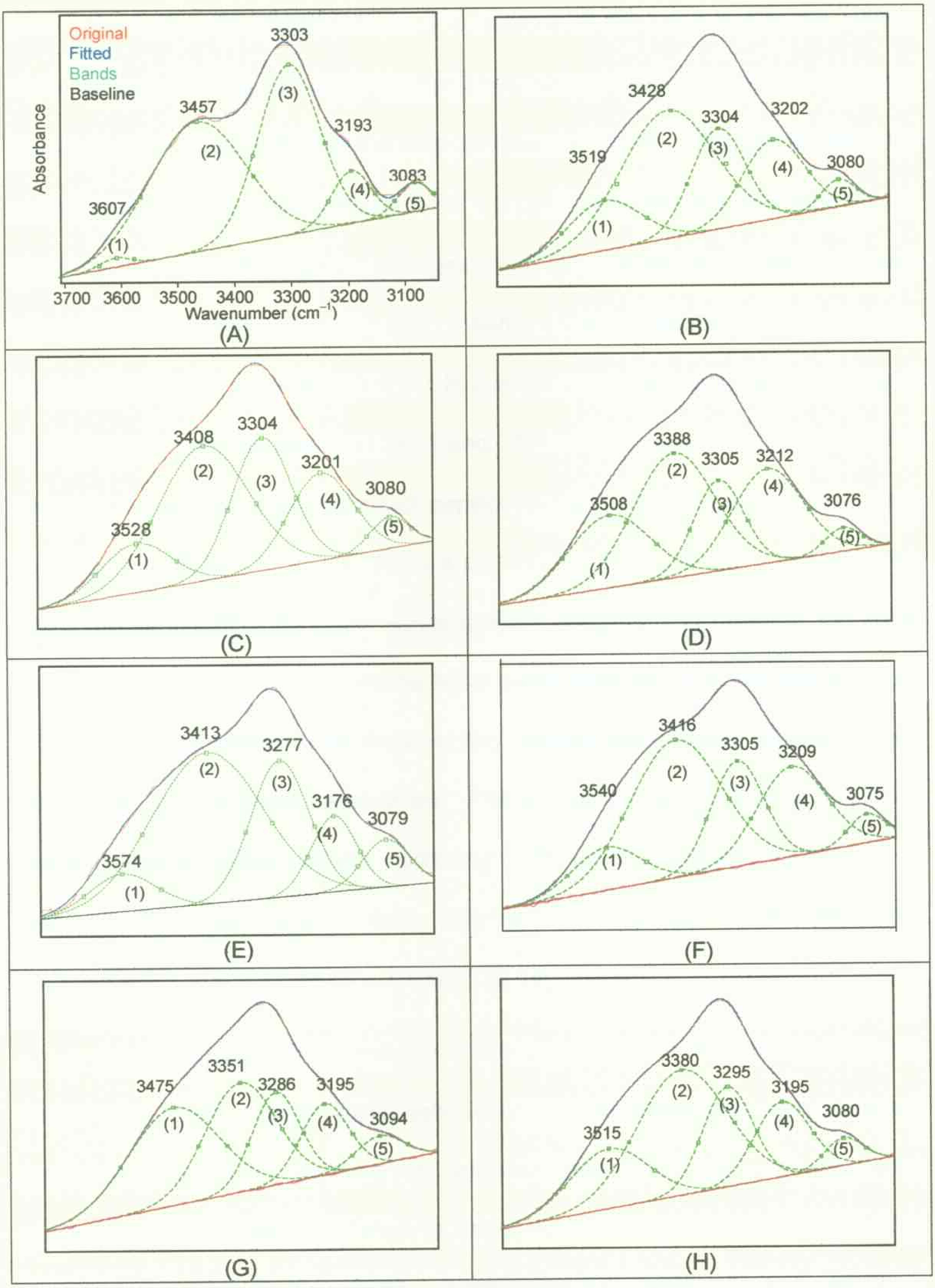


Figure 5.8. Curve-fitted FTIR spectra obtained from breast lesions; Region I of the spectrum given in Figure 5.7: (A) normal breast duct; (B) fibrocystic changes; (C) fibroadenoma; (D) intraduct papilloma; (E) DCIS; (F) LCIS; (G) IDC; and (H) ILC. This spectral region contains five major IR bands due to: (1) weakly hydrogen-bonded water ν(OH); (2) nucleic acids ν(NH_2); (3) amide A; (4) strongly hydrogen-bonded water ν(OH); and (5) amide B. Original trace: red; fitted trace: blue; fitted bands: green dotted lines; and baseline: brown.

Nucleic acid ν(NH_2) bands: The position of the band due to the ν(NH_2) mode of nucleic acids (3348–3459 cm^{-1}) has been used to distinguish between tissue components (epithelial and adipose tissues) [9,14]. The ν(NH_2) band for normal breast ducts was found at 3457 ± 2 cm^{-1}, whereas the positions of the bands for benign fibrocystic changes and fibroadenoma shifted to lower wavenumbers, 3428 ± 3 cm^{-1} and 3408 ± 4 cm^{-1}, respectively. Similar observations were made from the FTIR spectra from thin-sections of DCIS and LCIS biopsies with the peaks at 3413 and 3416 cm^{-1}, respectively. The ν(NH_2) bands in the spectra from intraduct papillomas, and IDC and ILC lesions were observed at 3388 ± 5 cm^{-1}, 3351 ± 4 cm^{-1}, and 3380 ± 4 cm^{-1}, respectively, which are at significantly lower wavenumbers compared to those in the FTIR spectra from normal tissue. Smaller band areas were found in the spectra from normal tissue compared with those in the spectra from benign lesions, *in-situ* carcinomas and invasive carcinomas. These findings on band positions were slightly lower in wavenumbers compared with those obtained from previous research by Gao and Eckel (Normal: 3434 ± 5 cm^{-1}, fibroadenoma: 3444 ± 5 cm^{-1}, and IDC: 3419 ± 10 cm^{-1}) [9,14].

Amide A band: The positions of amide A bands in the FTIR spectra are similar in normal, benign (fibrocystic changes, fibroadenoma and intraduct papilloma) and LCIS sections at approximately 3304 cm^{-1}, whereas the positions of the bands are shifted to lower wavenumbers in the spectra from DCIS (3277 ± 4 cm^{-1}), IDC (3286 ± 3 cm^{-1}) and ILC (3295 ± 1 cm^{-1}) sections. The spectral data from both the normal and fibroadenoma lesions are consistent with those found in previous research by Gao and Eckel, while this band in the spectra of an IDC lesion was observed at 3194 ± 5 cm^{-1} in the published work of Gao [9,14].

Amide B band: The lineshapes of the amide B band showed considerable variations in the breast tissue sections in both original and curve-fitted spectra (Figure 5.8), which indicated that there were different contributions from the amide B (protein) band within the specific area of the tissue. The position of the bands were centred at approximately 3080 ± 5 cm^{-1} for all tissue sections except in the spectrum of the IDC lesion, where the band shifts to higher wavenumbers (3094 ± 4 cm^{-1}).

Table 5.4. Curve-fitted spectral differences of representative FTIR spectra obtained from various breast lesions in Region I.

	Band Position (cm^{-1})				
Tissue Type (Samples)	**WBW (1)**	**Nucleic Acid (2)**	**Amide A (3)**	**SBW (4)**	**Amide B (5)**
(A) Normal (2)	3607±4	3457±2	3303±1	3192±2	3082±1
(B) Fibrocystic changes (20)	3519±8	3428±3	3304±2	3202±3	3080±2
(C) Fibroadenoma (4)	3528±6	3408±4	3304±2	3201±3	3080±4
(D) Intraduct Papilloma (3)	3508±2	3388±5	3305±2	3212±2	3076±4
(E) DCIS (5)	3574±4	3413±2	3277±4	3176±5	3079±3
(F) LCIS (2)	3540±6	3416±3	3305±3	3209±1	3075±2
(G) IDC (36)	3475±4	3351±4	3286±3	3196±3	3094±3
(H) ILC (11)	3515±7	3380±4	3295±1	3195±3	3080±3

5.2.2 REGION II: 3050-2750 CM^{-1}

The spectral region containing the bands due to CH stretching modes (depicted in Figure 5.9), shows significant variations in the band intensities and positions amongst all breast lesion. As described in Section 5.1, this region contains bands due to C–H vibrations, which originate from biomolecules and exogenous chemicals. The normal tissue spectrum exhibited strong ν(CH_2) bands and a weaker aromatic ν(CH) band at 3005 cm^{-1}, which are indicative of a high adipose content. The spectrum obtained from the fibroadenoma sample had a striking feature at 2960 cm^{-1}, which is attributed to the $\nu_{as}(CH_3)$ mode of lipids (Figure 5.9C), while a similar shoulder was also observed in the ILC spectrum (Figure 5.9H), which appeared as a weak shoulder on the $\nu_{as}(CH_2)$ band in the spectra of the other lesions. A strong $\nu_s(CH_2)$ band at 2980 cm^{-1} was also present in the fibroadenoma spectrum, which appeared as only a shoulder in the other breast lesion spectra. The shift of both $\nu(CH_2)$ bands observed in the DCIS spectrum (Figure 5.9E) was attributed to the presence of residual paraffin found in the deparaffinised FFPE section (Chapter 4, Section 4.4).

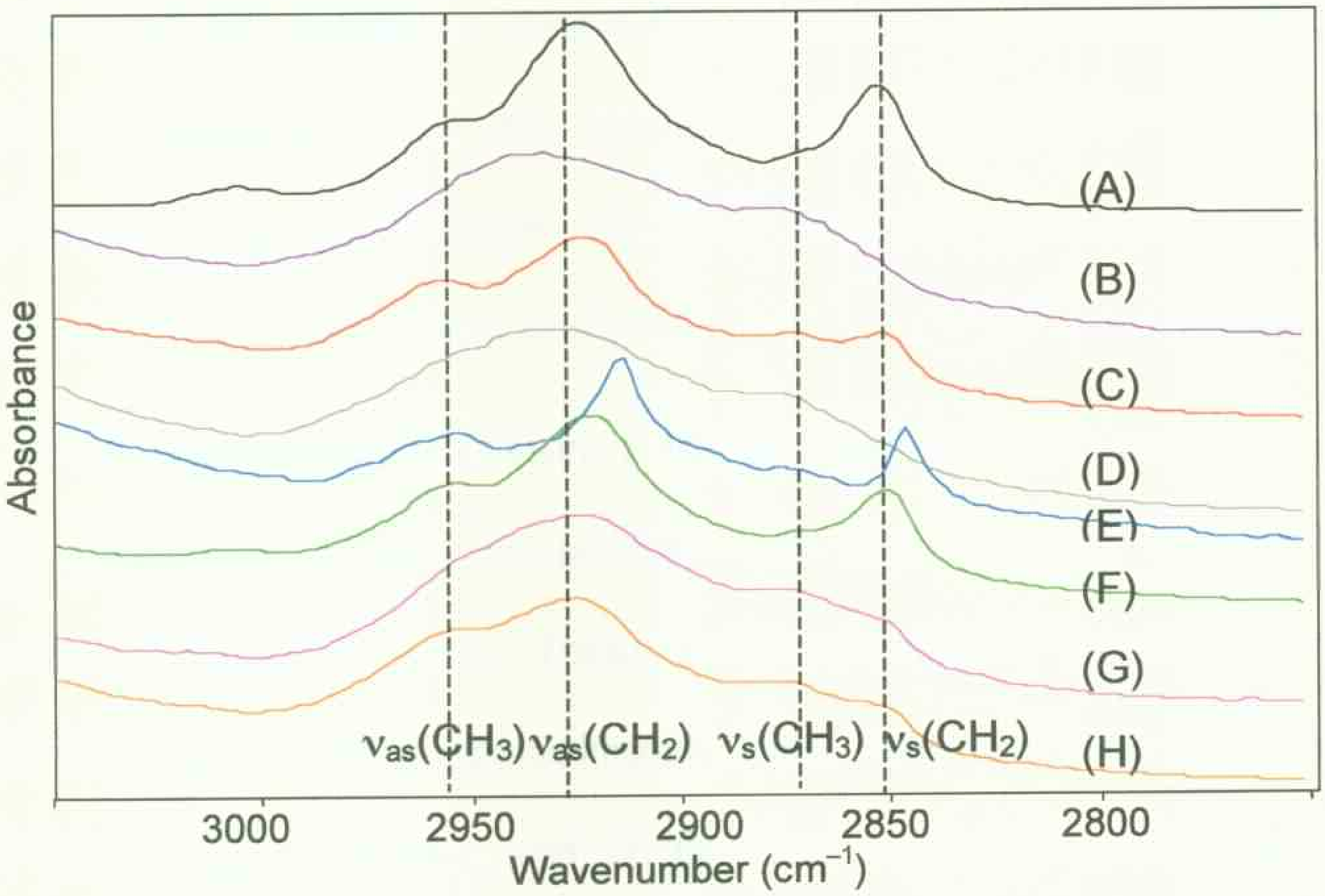


Figure 5.9. Expanded 3050-2750 cm^{-1} spectral region from the FTIR spectra in Figure 5.6: (A) normal breast duct; (B) fibrocystic changes; (C) fibroadenoma; (D) intraduct papilloma; (E) DCIS; (F) LCIS; (G) IDC and (H) ILC.

The results have demonstrated that the heterogeneous nature of breast tissue and the contribution of the residual paraffin have resulted in overlapping bands in the CH stretching region that makes this region unsuitable for routine tissue classification. Similar to Region I, this spectral region is not commonly used in diagnostic cluster analysis [3,9,21,26,38]. However, the spectral information in this region could be used to determine the percentage of adipose composition within the section.

5.2.3 REGION III: 1800–800 CM^{-1}

Figure 5.10 displays the bands in the fingerprint region (1800–800 cm^{-1}) of the spectra obtained from the tissue sections; this is the spectral region typically used to determine disease status [3,6,7,9,16,18,21,26,27,29,34,38-41]. The breast tissues were classified based on changes to the biological components: proteins (including collagens), lipids, and nucleic acids. A number of approaches were applied for diagnostic purposes including: visual examination of the spectra to detect the presence or absence of diagnostic bands; the evaluation of the second-derivative spectra and spectral comparisons using multivariate statistical analysis (principal component analysis and hierarchical cluster analysis).

In this spectral region (Figure 5.10), all tissue sections produced similar spectral information except cryosections with high adipose content (major component of female breast). These samples exhibited bands of strong to medium intensity centred at approximately 1745, 1465 and 1163 cm^{-1} assigned to the $\nu(C{=}O)_{\text{lipid-ester}}$, $\delta(CH_2)$, and $\nu(CO{-}O{-}C)$ modes of lipids, respectively [31,42] Other notable differences amongst the spectra from the lesions were the stronger absorptions of $\nu_{as}(PO_2^-)$ and $\nu_s(PO_2^-)$ modes of nucleic acids at approximately 1280 and 1098 cm^{-1}. The amide I band shifted from 1658 cm^{-1} in the spectrum from normal tissue to 1640 cm^{-1} in the spectrum from the DCIS section and the amide II band shifted from 1552 cm^{-1} (normal tissue) to 1536 cm^{-1} (DCIS). The 20 and 16 cm^{-1} amide I and II band shifts were due to the protein structural alternations during the formalin-fixation and subsequent paraffin-embedding, which corroborated the findings in Section 4.6 and those reported by Faolain and Pleshko [43,44].

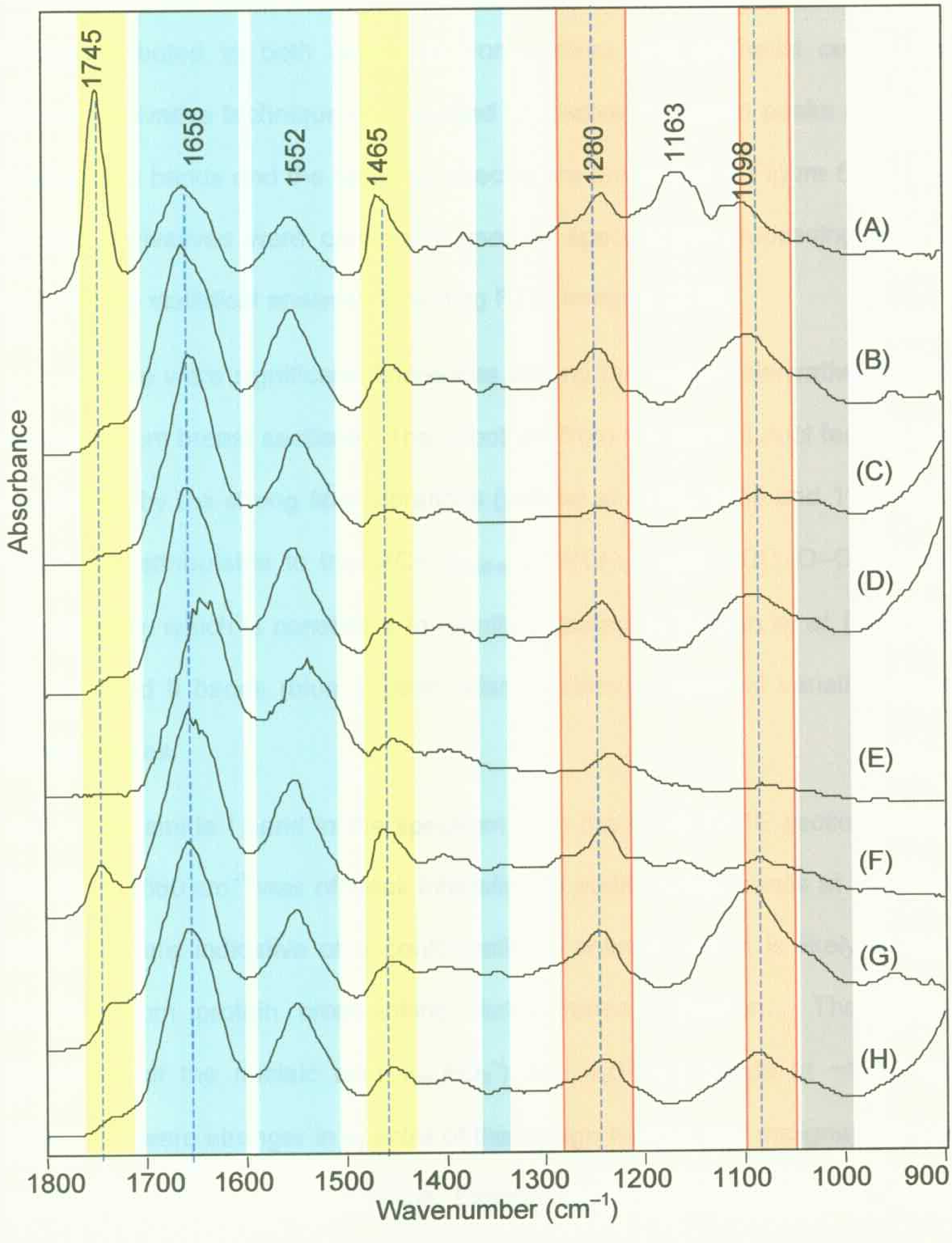


Figure 5.10. Expanded spectral fingerprint region from the FTIR spectra in Figure 5.6: (A) normal; (B) fibrocystic disease; (C) fibroadenoma; (D) intraduct papilloma; (E) DCIS; (F) LCIS; (G) IDC; and (H) ILC. Spectra are to scale and are offset for clarity. Areas of biological components marked in: (yellow) lipids; (blue) proteins; and (orange) nucleic acids. The grey region is used to classify disease states using hierarchical cluster analysis.

Visual examination of the IR spectra of breast lesions was not practical for reliable diagnostics, as these bands (proteins, lipids and nucleic acids) were overlapped with those from other biochemical components, e.g., the protein bands attributed to both collagens from stroma and epithelial cells. The second-derivative technique was applied to resolve the broad peaks under the overlapping bands and the resulting spectra are presented in Figure 5.11. The second derivatives were commonly used in spectral pre-processing prior to multivariate statistical analysis, including FTIR imaging [3,26].

There were significant differences among the second-derivative spectra obtained from breast sections. The spectrum from the normal duct feature was dominated by the strong lipid vibrations (yellow) at 1745, 1455 and 1163 cm^{-1}, which are attributable to the $\nu(C{=}O)_{\text{lipid-ester}}$, $\delta(CH_2)$, and $\nu(CO{-}O{-}C)$ modes, respectively, which is consistent to results obtained by Fabian *et al.* [42]. The amide I and II bands (blue, protein) also exhibited a spectral variation for all breast tissues.

The amide I band in the spectrum from the DCIS FFPE section (Figure 5.11E) at 1660 cm^{-1} was of weak intensity. In addition, two bands at 1645 and 1630 cm^{-1} are indicative of a conformational change, which is likely to have resulted from protein cross-linking during formalin fixation. The relative intensities of the nucleic acid $\nu_{as}(PO_2^-)$ and $\nu_s(PO_2^-)$ bands at ~1240 and 1080 cm^{-1} were stronger in spectra of the benign tumour and malignant tissues compared to those in normal tissue. Examination of the 1076-1038 cm^{-1} region (grey) revealed spectral differences attributed to variable glycoprotein mucin content [21].

To conclude, pre-processing of the original data is required to determine more accurate information. A combination of second-derivative and curve-fitting procedures can be used to reveal subtle variations in band positions, which could not be detected by visual examination of the original data.

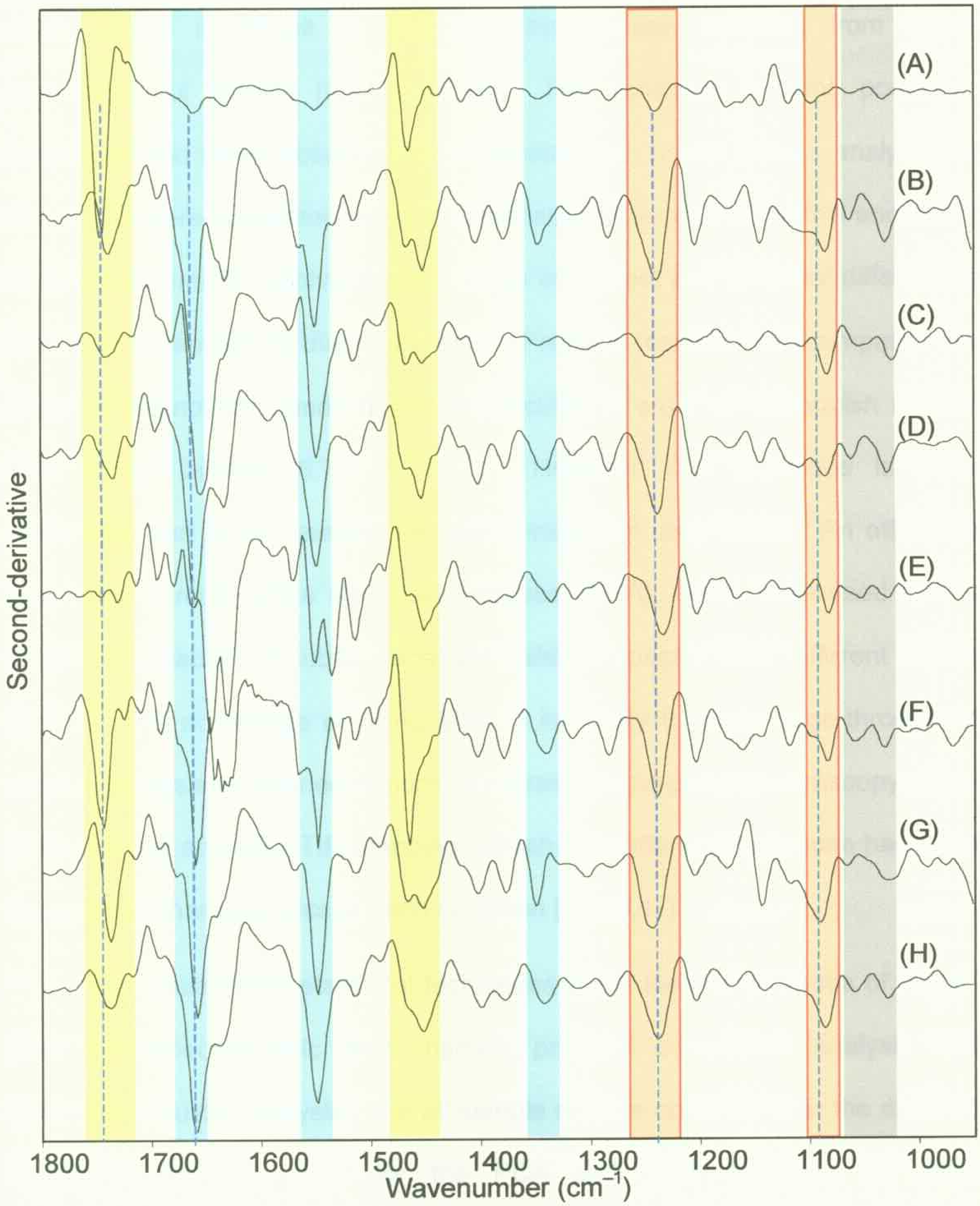


Figure 5.11. Second-derivatives of the spectra contained within Figure 5.10: (A) normal; (B) fibrocystic disease; (C) fibroadenoma; (D) intraduct papilloma; (E) DCIS; (F) LCIS; (G) IDC; and (H) ILC. Spectra are to scale and are offset for clarity. Areas of biological components marked in: (yellow) lipids; (blue) proteins; and (orange) nucleic acids. The grey region is used to classify disease states using hierarchical cluster analysis.

5.2.4 MULTIVARIATE STATISTICAL ANALYSIS OF BREAST TISSUES (SINGLE-POINT SPECTROSCOPY)

One of the challenges for tissue classification using single-point infrared microspectroscopy is the spectral similarity apparent in all biological samples, which makes it difficult to determine the disease condition from visual examination of spectra [9,14,15,42,45]. The variations in band positions, lineshapes and areas observed in the spectra from breast lesions analysed in this project were delineated using a combination of second-derivative and curve fitting procedures to resolve spectral areas of interest and to reveal differences that could be applied to diagnostic tests. Previous research has compared the areas of diagnostics bands from biomolecules in order to distinguish normal from benign lesions [5,7,9,14,15,17]. However, this technique is time-consuming and is not feasible for large-throughput diagnostics. An attractive alternative is multivariate statistical analyses (MSA), which were used in this project to establish if subtle variations exist in spectra from different breast lesions. The advantage of using MSA is its ability to apply large-throughput spectral analysis on defined regions of interest (single-point spectroscopy, FPA-FTIR imaging and SR-FTIR mapping), which can differentiate tissues based on the spectral characteristics of each condition [40,46,47].

Two multivariate statistical techniques were used for analyses of single-point microspectroscopic data, namely principal component analysis and hierarchical cluster analysis. Not all sample spectra collected over the duration of this study were used for the PCA and HCA analyses, as initial experimentation involved optimisation of the experimental protocols using a variety of sectioning techniques and mounting materials (Sections 4.6 and 4.7). Therefore, only the spectra collected using the same sectioning conditions, i.e., cryosectioning or FFPE sectioning, were analysed together. It was also essential to evaluate a small data set with known histopathological outcomes using the fingerprint region or specific peak regions in any MSA to understand where the major spectral variation(s) occurred amongst the sample spectra, prior to examination of a large data set containing various disease conditions,

5.2.4.1 PRINCIPAL COMPONENT ANALYSIS

Principal component analyses were applied to all of the IR spectra collected from breast tissues that had been subjected to the same sampling preparation and presentation techniques (sectioning and sample mounting). The analysis aimed to model the data with the least number of principal components. For example, if the adipose and epithelial tissues were compared, by examining the lipid and nucleic acid IR bands in the fingerprint region, the first two principal components (PC) should represent the major variance by separating the adipose and epithelial tissues into two distinguishable locations on the PC2 *vs.* PC1 scores plot.

Figure 5.12 presents the PC2 *vs.* PC1 scores plot from a data set containing six breast lesions including three benign (a fibrocystic change, a fibroadenoma, and an intraduct papilloma) and three malignant (IDC) lesions over the 1700–1000 cm^{-1} spectral region containing diagnostic bands of proteins, nucleic acids, adipose and stromal tissues. These samples were cryosectioned to a thickness of 5 μm and the sections were mounted on CaF_2 windows. Previously, PCA using this spectral region was used to distinguish between normal and different benign and malignant tumour types [3,21,26,42].

PC1 describes the major differences (94% of total variance) between the spectra of the benign fibrocystic changes and other lesions, while PC2 (5.5%)

describes the differences between the benign fibroadenoma, intraduct papilloma and the IDC lesions. From this small test data set, PCA has provided indications that it will be able to separate different groups of lesions on a larger data set based on the FTIR spectra in the fingerprint region (1700–1000 cm^{-1}) since the first two PCs have over 99.5% of total variance. However, both fibroadenoma and intraduct papilloma spectra were unable to be separated, which may be due to the similarity of the epithelial cellular features (non-cancerous) from the samples that were examined.

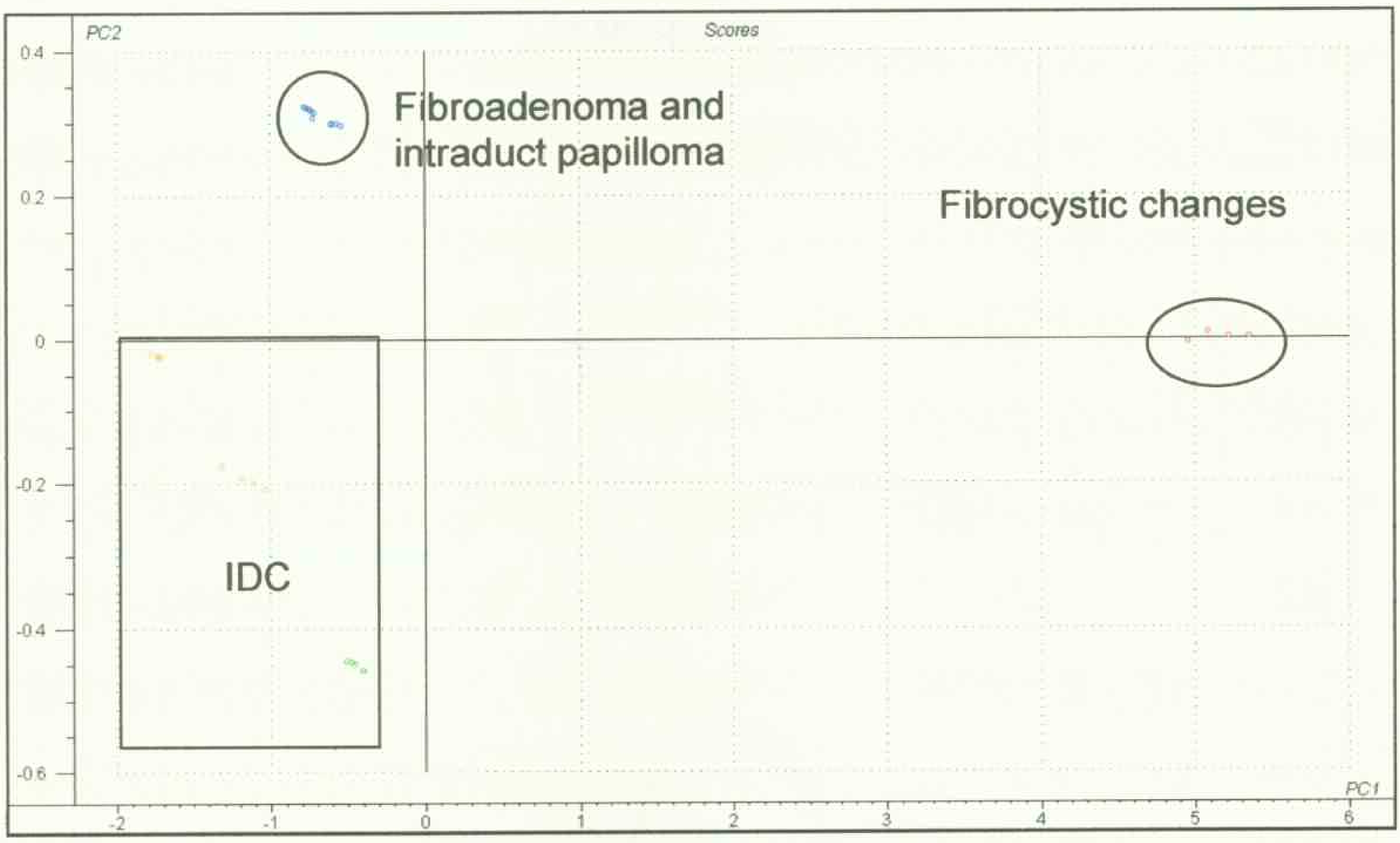


Figure 5.12. Scores plots for PC2 *vs.* PC1 obtained from a training data set that contained spectra from three benign and three malignant lesions in the spectral region of 1700–1000 cm^{-1}. Experimental conditions: 5-μm-thick cryosections mounted on CaF$_2$ windows, transmission; 512 scans; 4 cm^{-1}; ×15 magnification. Each point represents an IR spectrum collected from the tissue surrounding the duct.

As mentioned in Sections 5.2.1 and 5.2.2, spectral regions I and II contained bands that are not commonly used for routine diagnostics amongst cancerous studies [3,21,26,42]. Consequently, the entire mid-infrared spectrum is not commonly used in MSA in this research. In addition, tissues used in this data set were deparaffinised FFPE sections mounted on transflective slides that would presumably be affected by the residual paraffin and anomalous dispersion artefacts (Section 4.8). As a result, most of this research and other published research use the fingerprint region for MSA including PCA and HCA [3,21,26,42].

A larger sample size of FFPE breast sections (52 samples), which consisted of 15 fibrocystic changes; three fibroadenomas; two intraduct papillomas; five DCIS; two LCIS; 15 IDC and 10 ILC, was analysed using PCA over the 1700–1000 cm^{-1} spectral region. The PCA results using the raw data and the second-derivatives of these spectra are presented in Figures 5.13 and 5.14, respectively.

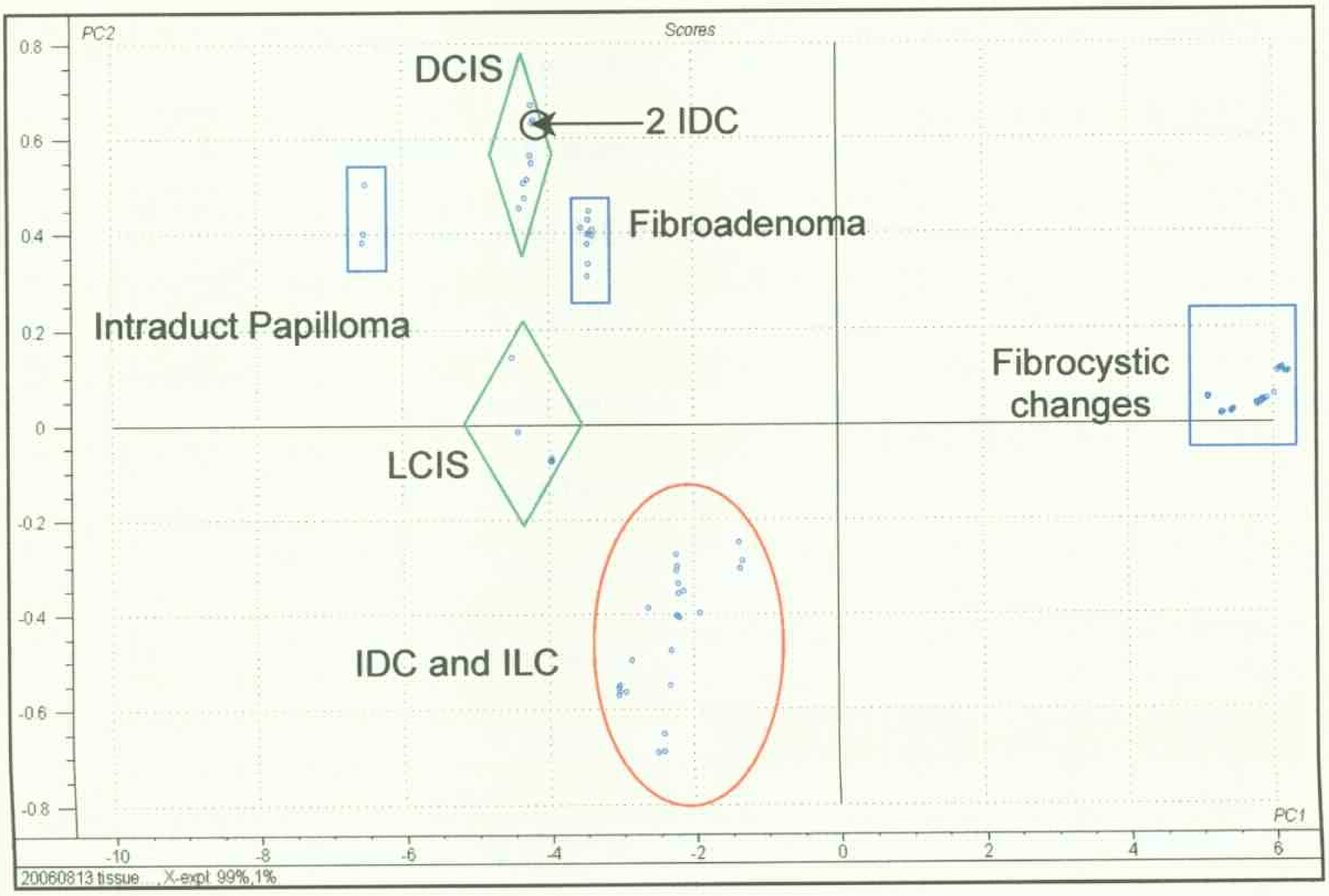


Figure 5.13. Scores plot (PC2 *vs.* PC1) for the differentiation of FTIR spectra that show a separation of the major disease conditions by the use of the information in the 1700–1000 cm^{-1} spectral region. Samples were 5-μm-thick mounted on transflective slides: 15 fibrocystic changes; 3 fibroadenoma; 2 intraduct papillomas; 5 DCIS; 2 LCIS; 15 IDC; and 10 ILC. Experimental conditions: transflection mode; 512 scans; 4 cm^{-1}; ×15 magnification. Each point represents an average of triplicate measuring in the major pathological features using the adjacent H&E section. More than one section from each sample were measured.

As was the case in Figure 5.12, PC1 in Figure 5.13 accounts for 99% of total spectral variance, which clearly distinguishes the IR spectra from samples that contain fibrocystic changes from those obtained from other breast lesions. PC2 accounts for the remaining 1% of the total spectral variance, which separates the spectra from benign (fibroadenoma and intraduct papilloma) sections and *in-situ* carcinomas (mainly DCIS) from the malignant IDC and ILC lesions. The spectra from LCIS lesions were classified as being intermediate between the spectra obtained from benign and malignant lesions, as expected from the progression of the disease. Although lesions involving fibrocystic changes are sometimes classified as benign, the morphological manifestation is enlargement of the breast duct without any alternations in cellular compositions [48]. In contrast, both fibroadenoma and intraduct papilloma involve epithelial proliferation and, therefore, the spectra were clustered with other diseased lesions. The difference between DCIS and LCIS lesions is that the latter are rare and often co-exist with a ILC lesion, which makes LCIS occult to histopathology [49]. There were two samples diagnosed by histopathology as malignant IDC lesions were classified by IR/PCA as DCIS. This could be due to the co-existence of both DCIS and IDC features on the same malignant IDC lesions. These two IDC spectra were measured at the regions associated with DCIS, which clustered with the DCIS group.

The PC analysis using the raw data could not distinguish between the IDC and the ILC lesions, but a PCA scores plot based on the second-derivatives of the spectra (Figure 5.14) separated the benign lesions from the *in-situ* carcinomas and invasive carcinomas in the PC1 direction (74% of total spectral variance). The second derivatives of these spectra separated the IDC and DCIS sections from the ILC and LCIS in the PC2 direction (14%). The DCIS spectra were closely related to the IDC spectra, whereas there was a clear difference separating the ILC and LCIS spectra, which was not previously reported. As a result, the second-derivative PCA scores plot using the fingerprint region differentiated benign lesions and carcinomas (both *in-situ* carcinomas and invasive carcinomas) with high sensitivity (100%). Two IDC lesions were grouped together with the DCIS lesions, which were consistent with the PCA result previously shown (Figure 5.13). The subtle spectral variations identified by using the second-derivative PCA over the fingerprint region could improve the current histopathological diagnosis for the rare LCIS condition [49].

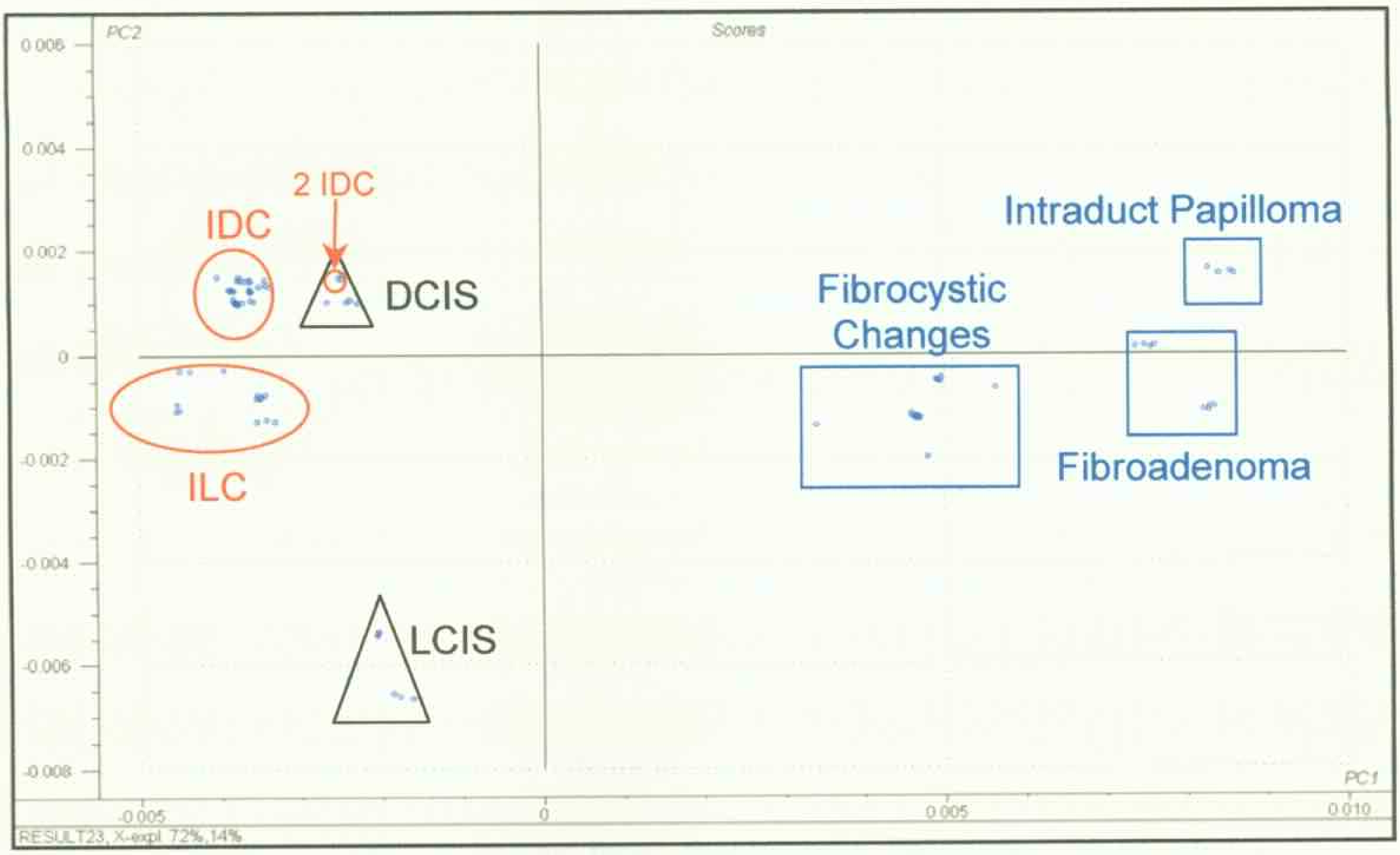


Figure 5.14. Scores plot (PC2 *vs.* PC1) for the differentiation of FTIR spectra that show a separation of the major disease conditions by the use of the information from the 1700–1000 cm^{-1} spectral region, displayed as nine-point Savitzky-Golay second-derivatives of the spectra contained within Figure 5.13.

5.2.4.2 HIERARCHICAL CLUSTER ANALYSIS

As with principal component analysis, it is important to begin the hierarchical cluster analysis (HCA) with a small spectral data set containing known samples, and to understand the fundamental changes of the 'signature' peaks in the spectra prior to interpretation of any large HCA data set. A HCA (Figure 5.15) was performed on a data set consisting of spectra from ten 5-μm-thick benign cryosections mounted on transflective slides, including two fibrocystic changes (A), one fibroadenoma (B), one intraduct papilloma (C), and four IDC breast lesions (D). Two samples (X and Y) without prior knowledge of disease states were also classified by HCA.

The analysis was performed by comparing the second-derivatives of the spectra using Ward's clustering algorithm in the 1720–1487 cm^{-1} region, specifically to examine the spectral variations in the amide I and II bands. The amide I band is sensitive to conformational changes in protein secondary structures and may be used to differentiate breast conditions [3]. Infrared spectra of all lesions were collected from samples that were cryosectioned, mounted on CaF_2 windows and were measured at the characteristic area of each disease type.

There are two broad classification groups obtained from this data set: fibrocystic changes (A) and a group including fibroadenoma (B) and intraduct papilloma (C) and IDC (D). The HCA result (Figure 5.15) correlated with the PC1 results, where spectral differences separated spectra obtained from tissues with fibrocystic changes from those from fibroadenomas and intraduct papillomas. However, examination of the sub-groups revealed that one fibroadenoma (B) sample clustered with one IDC lesion, while an intraduct papilloma spectrum clustered with the remaining IDC samples. This may be due to the small data set used in this preliminary test. The HCA clustered the spectra of the two unknown samples with those from tissues exhibiting fibrocystic changes. These unknown breast tissue sections were subjected to histopathology after their spectroscopic analyses and were both diagnosed as having fibrocystic changes. Although the data set for this initial test was small and not all sub groups clustered correctly, HCA successfully identified the two unknown samples as belonging to fibrocystic changes. These results were confirmed by histopathology with 100% accuracy.

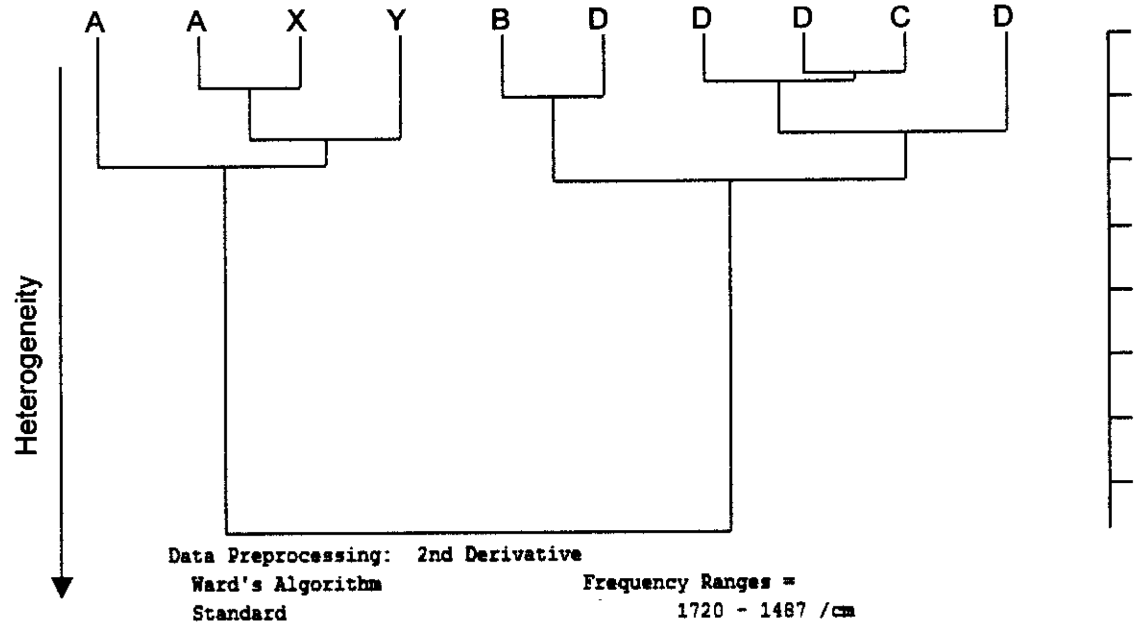


Figure 5.15. HCA classification of various benign breast cryosections using second-derivative data preprocessing, and Ward's clustering algorithm (1720–1487 cm^{-1} spectral region). A – Benign fibrocystic changes (2); B – fibroadenoma (1); C – intraduct papilloma (1); D – invasive ductal carcinoma (4); X and Y –samples not previously diagnosed. Numbers of samples are marked in brackets.

A large-sample HCA was performed from the same data set used to produce the results in Figure 5.14: 15 fibrocystic changes, three fibroadenoma, two intraduct papilloma, five DCIS, two LCIS, 15 IDC, ten ILC, and two unknown samples (X and Y). As discussed in Section 5.2, spectral variations were observed, particularly in Region III, in all FTIR spectra obtained from different breast lesions, although this spectral region is not commonly used for spectroscopic tissue differentiation. However, it could provide additional information to improve the disease classification.

The second-derivatives of the spectra obtained from this large sample dataset were vector-normalised and Ward's clustering algorithm was applied to the following regions: 3705–3001 cm^{-1}, 1716–1585 cm^{-1}, and 1140–1038 cm^{-1}. The region under analysis included information about the hydrogen-bonding of residual water in the tissues, $\nu(NH)$ of proteins and $\nu(NH_2)$ of nucleic acids, the protein secondary structures (amide I) and the $\nu_s(PO_2^-)$ of nucleic acids, which was found to produce better classification using a wider spectral range [10,42]. Figures 5.16 and 5.17 show the HCA results in the above regions and the $\nu_s(PO_2^-)$ region, respectively.

The dendrogram, as shown in Figure 5.16, classified two major tissue groups, benign (A to C) and malignant (F and G). Although DCIS (D) and LCIS (E) spectra were not classified as invasive (malignant), they clustered within the malignant group. As the heterogeneity increased, more spectral variations were observed amongst the FTIR spectra obtained from breast lesions. Spectra obtained from DCIS and LCIS sections have been classified as malignant, however, within the sub group, they are well-separated from the IDC and ILC lesions, as evident by the heterogeneity value of the malignant group. This may be due to the transitional nature of these conditions and the presence of cancerous cells within the ducts or lobules of the breast, which have more spectral similarities than those classified as benign. Similarly, IDC and ILC spectra were separated successfully using the combined spectral region with a heterogeneity value of ~2.75. Two samples diagnosed by histopathology as DCIS clustered within the IDC sub-group (F), which were not separated in the PCA score plots in Figures 5.13 and 5.14. This could be due to the IR spectra were collected at the area associated with DCIS co-existed within the unstained malignant IDC lesions. The IR spectra of two pre-classified malignant IDC tissue sections were clustered within the DCIS (D) sub-group, which were consistent with the PCA results. These two lesions were subsequently stained after IR analysis and histopathology results confirmed that these lesions were, in fact, high-grade DCIS in which cell proliferation was in the process of invading through the basement membrane, as evident from the adjacent H&E stained section. The HCA classification and subsequent histopathology examination overruled the original classification of being malignant IDC. The unknown samples (X and Y) were again clustered into the fibrocystic changes, which further confirm the tissue conditions. The HCA result is consistent with the PCA results discussed in the previous section.

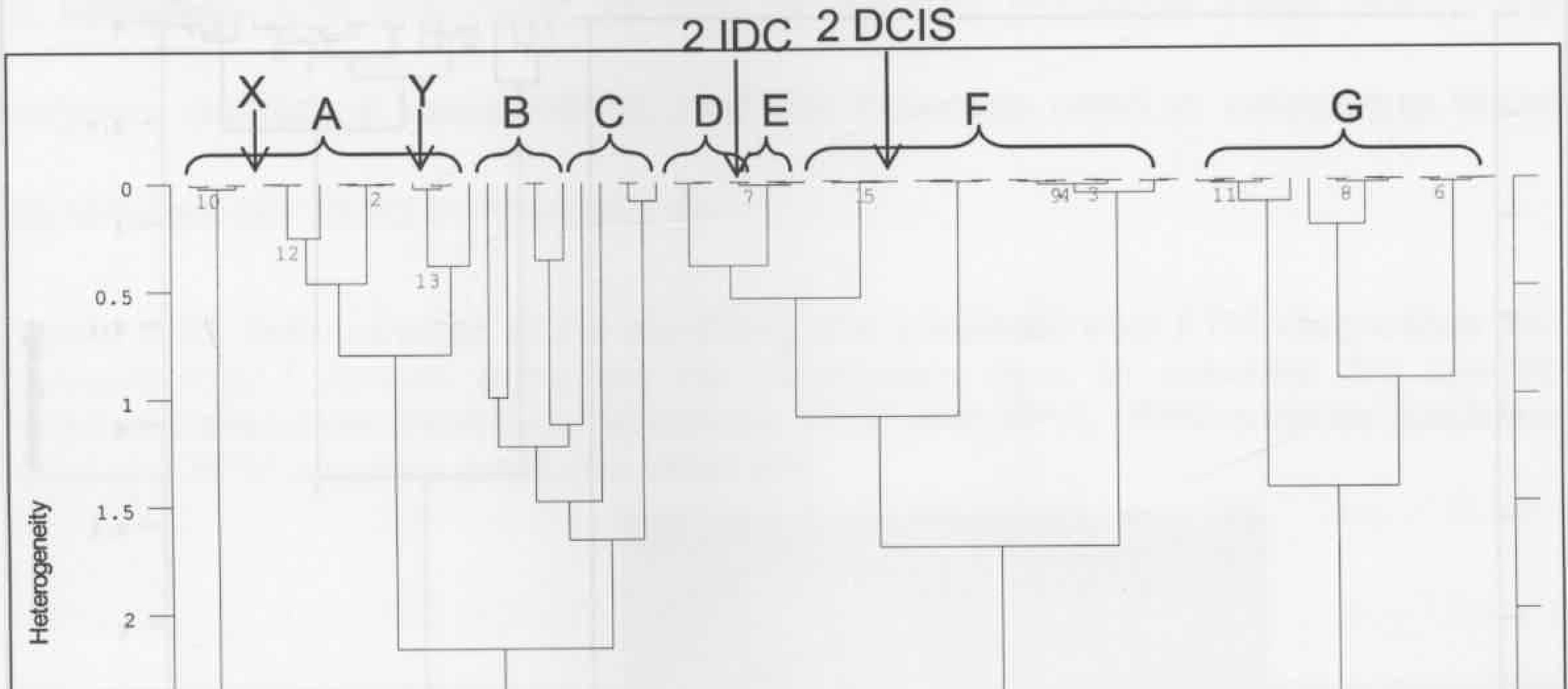


Figure 5.16. HCA plot separating breast lesions. Second-derivative spectra were vector-normalised, Ward's clustering algorithm were calculated in spectral regions of 3705–3001, 1716–1585 and 1140–1038 cm^{-1}. (A) fibrocystic changes; (B) fibroadenoma; (C) intraduct papilloma; (D) DCIS; (E) LCIS; (F) IDC and (G) ILC. (X) and (Y): samples not previously diagnosed.

To illustrate the advantage of using a wider spectral range compared with a specific band, the HCA dendrogram of the $\nu_s(PO_2^-)$ band attributed to nucleic acids was plotted (Figure 5.17). The dendrogram in Figure 5.17 clearly differentiates the spectra from benign lesions and from those obtained from the *in-situ* carcinomas and malignant lesions. A similar approach to differentiate colorectal adenocarcinomas was reported by Lasch *et al.* [21]. As expected, the analysis was unable to separate successfully the spectra obtained from *in-situ* carcinomas and malignant lesions. This is likely to be due to the increased cell proliferation found only within the ducts/lobules in *in-situ* carcinomas, while cellular invasion to surrounding breast tissues in invasive carcinomas.

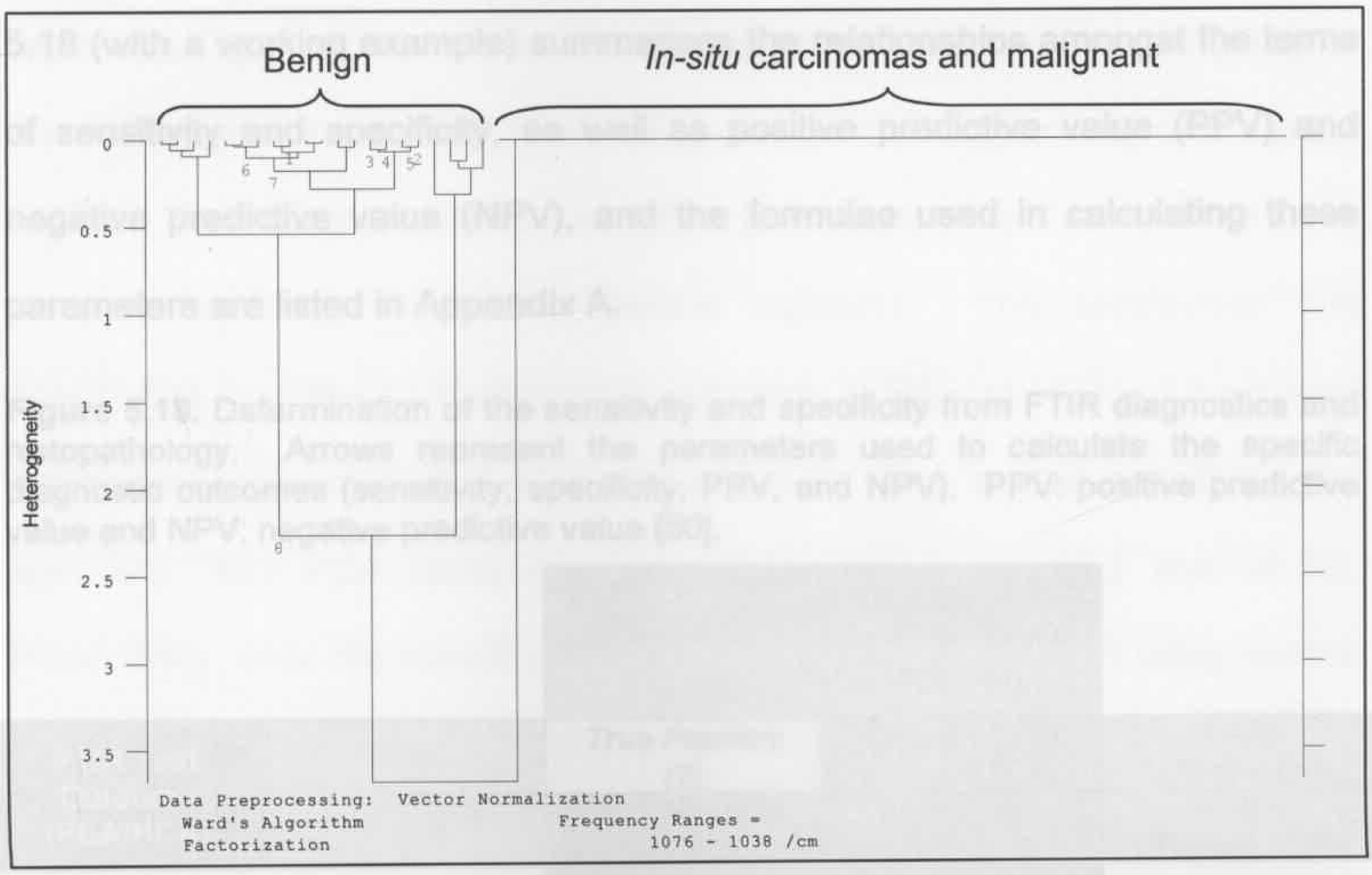


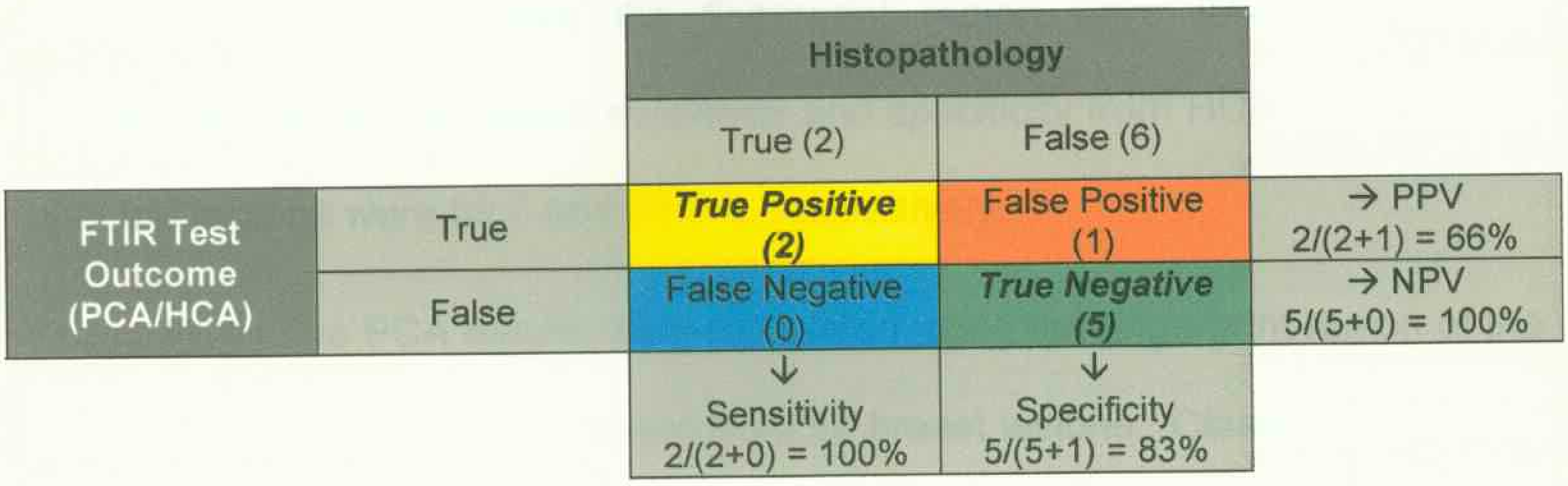
Figure 5.17. HCA plot of $\nu_s(PO_2^-)$ of nucleic acids separating benign and malignant lesions using the data from the 1078–1038 cm^{-1} spectral region. Second-derivative spectra were vector-normalised, D-values and Ward's clustering algorithm were calculated.

Similar to PCA, HCA has been demonstrated to be an alternative method for tissue differentiation by analysing the second-derivatives of the single-point spectra within the fingerprint region and/or specific band areas. Although the spectra from the *in-situ* carcinomas and malignant lesions were not well-separated using the 1078–1038 cm^{-1} spectral region, HCA could distinguish between the benign and carcinoma spectra with high heterogeneities (~3.5) in Figures 5.16 and 5.17.

5.2.4.3 SENSITIVITY AND SPECIFICITY OF PCA AND HCA OUTCOMES

Sensitivity and specificity are common statistical measures of how well a binary classification correctly identifies a positive and negative medical condition, respectively [50]. They are calculated based on the results from the FTIR test outcomes (PCA/HCA) and the gold standard histopathology. Figure 5.18 (with a working example) summarises the relationships amongst the terms of sensitivity and specificity, as well as positive predictive value (PPV) and negative predictive value (NPV), and the formulae used in calculating these parameters are listed in Appendix A.

Figure 5.18. Determination of the sensitivity and specificity from FTIR diagnostics and histopathology. Arrows represent the parameters used to calculate the specific diagnostic outcomes (sensitivity, specificity, PPV, and NPV). PPV: positive predictive value and NPV: negative predictive value [50].



		Histopathology		
		True (2)	False (6)	
FTIR Test Outcome (PCA/HCA)	True	***True Positive* (2)**	False Positive (1)	→ PPV 2/(2+1) = 66%
	False	False Negative (0)	***True Negative* (5)**	→ NPV 5/(5+0) = 100%
		↓ Sensitivity 2/(2+0) = 100%	↓ Specificity 5/(5+1) = 83%	

Sensitivity is calculated based on the true histopathologically classified condition and is compared with the FTIR test outcome including both true positives and false negatives. Specificity is determined by using the true negatives (false results in FTIR and histopathology) over the sum of true negative and false positive. Furthermore, PPV and NPV are determined using the true and false FTIR outcomes, respectively, regardless of the histopathology results.

Previous research by the IMRR group using ^{1}H MRS distinguished malignant tissues from benign lesions with a sensitivity and specificity of 95 and 96%, respectively [51]. Jackson, *et al.* reported that FTIR spectra obtained from low-grade, intermediate-grade, and high-grade breast tumours using cluster analysis within the fingerprint region achieved sensitivities of 91, 76.5 and 100% and specificities of 100, 97.6 and 85.9%, respectively [19].

The PCA (fingerprint region, Figure 5.14) and HCA (various spectral regions in Figure 5.16) results clearly differentiated the second-derivatives of the FTIR spectra obtained from malignant, *in-situ* carcinomas and benign lesions. Theoretically, both PCA and HCA techniques should achieve the same classification using the same spectral region(s). The sensitivities and specificities from the FTIR diagnostic outcomes (PCA and HCA) of the breast lesions are presented in Tables 5.5 and 5.6. The overall sensitivity and specificity from PCA using the fingerprint region were 97.3 and 98.5%, respectively, while the overall sensitivity and specificity from HCA using various spectral regions were 89.7 and 96.6%, respectively.

<u>PCA:</u> The PCA results were calculated using the fingerprint region of the second-derivative spectra obtained from 52 breast lesions. Classification of the benign lesions (fibrocystic changes, fibroadenomas and intraduct papillomas) was achieved with 100% sensitivity and specificity, as well as for PPV and NPV. Two malignant IDC sections were classified by PCA as *in-situ* carcinomas, the unstained IDC lesions were stained and examined by histopathologist after

FTIR analyses that these two lesions were classified as low-grade malignant IDC with DCIS morphological architecture. The sensitivity and specificity of the technique to diagnose *in-situ* carcinomas was 100 and 95.6%, respectively. The PPV, based on the true and false positives from the FTIR PCA outcomes, was 77.8% and since there were no false negative results, the NPV for *in-situ* carcinomas was 100%. 23 of the 25 samples that were histopathologically diagnosed as malignant were correctly classified using PCA. Two samples were identified as *in-situ* carcinomas, which yielded a sensitivity and specificity of 92 and 100% for malignant lesions, respectively. There were no false positive results but two false negatives were found, the PPV and NPV of the malignant lesions were 100 and 93.1%, respectively.

<u>HCA:</u> Two lesions without prior knowledge of histopathology results marked as unknowns, were classified as benign fibrocystic changes by the hierarchical cluster analysis (Figures 5.15 and 5.16). As a result, benign lesions were classified with a sensitivity and specificity of 100 and 94.1%, respectively. From the HCA result using various spectral regions, two samples histopathologically diagnosed as *in-situ* carcinomas were classified within the malignant group, and, two pre-diagnosed as malignant were classified by HCA as *in-situ* carcinomas, which were consistent with the PCA results (containing both IDC and DCIS features on the tissue sections). This resulted in a sensitivity and specificity for *in-situ* carcinomas of 77.8 and 95.6%, respectively, while the PPV and NPV were 77.8 and 95.6 per cent, respectively. The lower sensitivity and PPV outcome is proposed to be due to the small number of *in-situ* carcinoma samples examined in this research. Finally, the sensitivity and specificity for malignant lesions was 91.3 and 100%, respectively. The NPV of the malignant lesions was 93.9%, which was attributed to the samples being diagnosed as malignant by histopathology, but classified as *in-situ* carcinomas using HCA.

The combined spectroscopic results and MSA demonstrated that there is a high accuracy on discrimination between spectra obtained from benign and malignant lesions. The sensitivity and specificity from the FTIR MSA outcomes compared well with those spectra obtained using ^{1}H MRS [51]. Although there were discrepancies on diagnosis of *in-situ* carcinomas and malignant lesions from the results of both PCA and HCA, the gold standard histopathology was used as the finalised diagnosis. Alternative techniques to ensure the accuracy of the diagnosis could be performed using ^{1}H MRS (but this was not done in this research).

Table 5.5. Comparison of tissue diagnostic outcomes from the PCA results using the second-derivatives of FTIR spectra of breast lesions with those obtained from histopathology. PPV – positive predictive value; and NPV – negative predictive value. True positives (yellow); false positives (orange); false negatives (blue); and true negatives (green).

			Histopathology									
			Benign		*In-situ* Carcinoma		Malignant		Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
			20		*7*		*25*		n/a	n/a	n/a	n/a
			+	–	+	–	+	–	n/a	n/a	n/a	n/a
FTIR (PCA) Fingerprint Region	Benign	+	20	0	0		0		100	100	100	100
		–	0	32								
	In-situ Carcinoma	+	0		7	2	2		100	95.6	77.8	100
		–			0	43						
	Malignant	+	0		0		23	0	92	100	100	93.1
		–					2	27				
	Overall								**97.3**	**98.5**	**92.6**	**97.7**

Table 5.6. Comparison of tissue diagnostic outcomes from the HCA results using the second-derivatives of FTIR spectra of breast lesions with those obtained from histopathology. (*) Two samples without prior knowledge of histopathology results are marked as unknowns. PPV – positive predictive value; and NPV – negative predictive value. True positives (yellow); false positives (orange); false negatives (blue); and true negatives (green).

			Histopathology										
			Benign		*In-situ* Carcinoma		Malignant		Unknowns*	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
			20		*7*		*25*		*2*	n/a	n/a	n/a	n/a
			+	–	+	–	+	–	n/a	n/a	n/a	n/a	n/a
FTIR (HCA) Multi-Spectral Regions	Benign	+	20	2	0		0		2	100	94.1	90.9	100
		–	0	32									
	In-situ Carcinoma	+	0		7	2	2		0	77.8	95.6	77.8	95.6
		–			2	43							
	Malignant	+	0		2		21	0	0	91.3	100	100	93.9
		–					2	31					
	Overall									**89.7**	**96.6**	**89.6**	**96.5**

5.3 CHARACTERISATION OF BREAST TISSUE USING FPA-FTIR IMAGING AND MULTIVARIATE STATISTICAL ANALYSES

The third aim of this research was to characterise histopathological well-defined breast ductal diseases using focal-plane array (FPA) FTIR imaging together with multivariate statistical analysis (FCM, KMC, PCA, and HCA). This combination is the most powerful technique for tissue characterisation that could be used as a routine diagnosis for histopathology. The advantage of FPA-FTIR imaging has been described in Section 1.6.3.4. IR functional group maps generated from images collected from a normal breast duct, a DCIS and an IDC breast lesion together with multivariate statistical analyses are presented in the following section. All FFPE tissue sections were deparaffinised using xylene and/or histopure to minimise possible tissue contamination and the imaging data were vector-normalised prior to any statistical analysis. Each IR image has a dimension of 350 × 350 μm^2 (4096 spectra from a 64 x 64 FPA) and each pixel corresponds to an area of 6.25 × 6.25 μm^2.

5.3.1 NORMAL BREAST DUCT

Figure 5.19A presents the optical image of an unstained area of a breast tissue FFPE section with the duct being the dominant feature surrounded by adipose (white, dissolved) and stromal tissues. The box in the optical image shows the area imaged (350 x 350 μm^2, Figure 5.19B) with IR microspectroscopy and points (i) to (v) correspond to where individual single-point spectra (Figures 5.19C and 5.20) were extracted from the IR image. The duct lumen shown here was classified as normal by histopathology but the entire section was classified as a low-grade ductal carcinoma *in-situ*. This is due to the presence of tumour cells arranged in a cribriform architecture found within other ducts that are not shown in the optical image. As illustrated in Chapter 3, Figure 3.10 (analysed in subsequent sections), necrosis and micro-invasion were absent in this low-grade cribriform DCIS lesion [52].

As previously stated, cell proliferation results in increased levels of nuclei, proteins and lipids (cell membranes), with the most easily detected spectroscopic changes resulting from increased levels of nucleic acids [5,9,38]. Functional group maps are commonly used to illustrate the distribution of the diagnostic bands (proteins, lipids and nucleic acids) prior to any cluster analysis [3,21,42].

The functional group map shown in Figure 5.19B is based on the 1300–1000 cm^{-1} spectral region, which is mainly attributable to the nucleic acid $\nu_{as}(PO_2^-)$ and $\nu_s(PO_2^-)$ bands observed at 1240 and 1082 cm^{-1} but also has minor contributions from glycogen (carbohydrate) bands [8,31,42]. The marked areas (shown in dark red) are indicative of high levels of nucleic acids and low levels are depicted in dark blue. There is a high density of cells/cell nuclei at the locations coloured red than within the area of the duct lumen (blue). The dark blue areas found within the duct lumen and where the adipose tissues were, prior to being dissolved during the FFPE sectioning process (marked in white in A), are indicative of relatively low levels of nucleic acids in these areas, as expected.

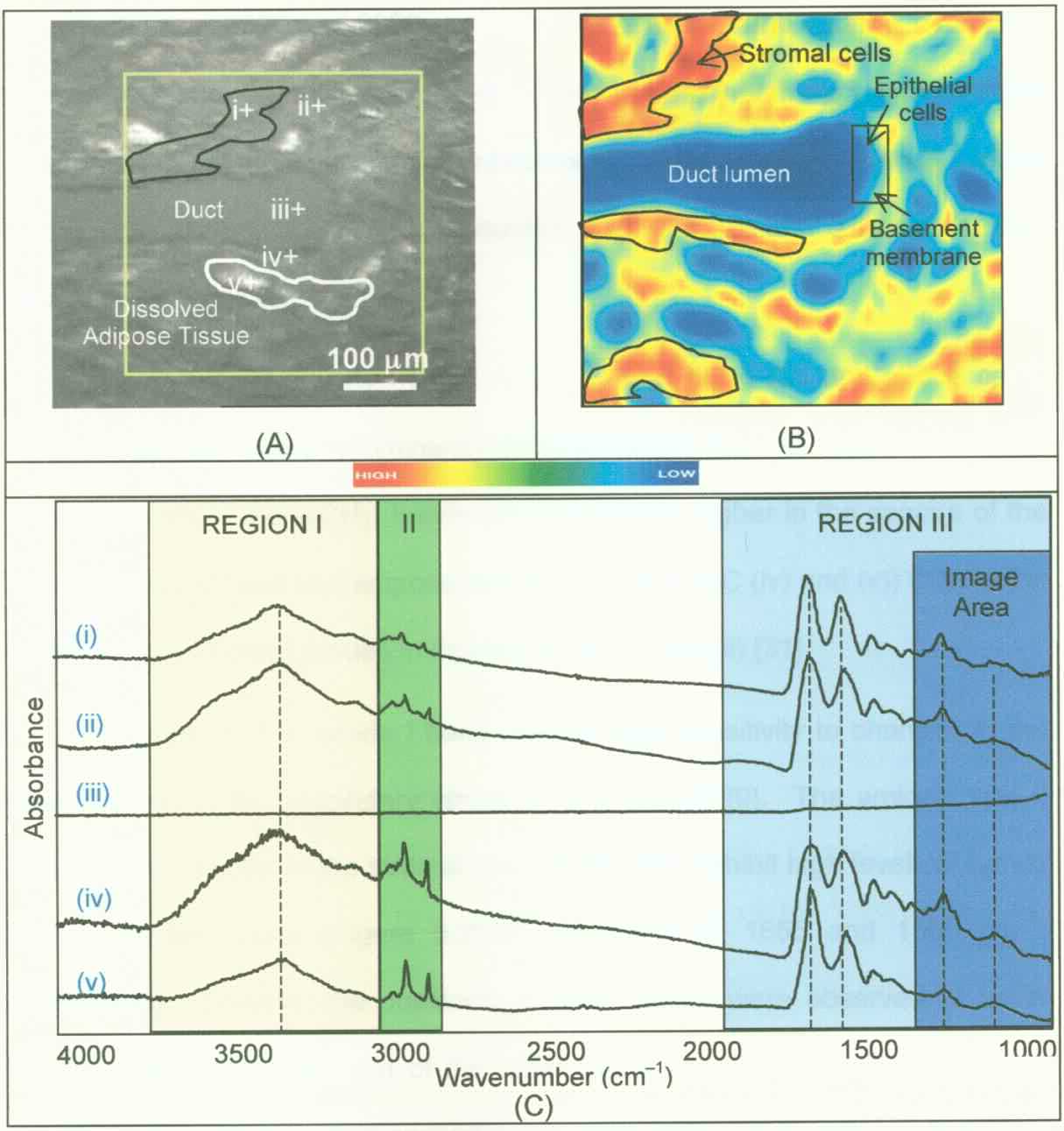


Figure 5.19. (A) White light image of an unstained 5-μm-thick breast FFPE section showing a 'healthy' breast duct and the area selected for infrared imaging is marked by the black square, (B) false-colour IR image of the 1300–1000 cm^{-1} region taking in the region marked in (A); (C) normalised IR spectra, locations of extracted spectra are marked (+): (i) stroma exhibiting high nucleic acid levels; (ii) stroma; (iii) duct lumen; (iv) basement membranes of ducts; and (v) adipose tissue. Experimental conditions: 64×64 array; 256 scans; 2 cm^{-1} res.; ×15 magnification; image area: 350 × 350 μm^2.

Representative infrared spectra of individual breast components are presented in Figure 5.19C and have been normalised to the amide I band centred at ~1650 cm^{-1}. Interpretation of the spectra was approached by dividing the spectrum into three regions as described in Section 5.1.

Region I: The amide A and B bands at 3380 and 3080 cm^{-1}, along with the shoulders at 3520 and 3200 cm^{-1}, arise from weakly and strongly hydrogen-bonded ν(O–H) modes of water [9,14]. These are more dominant in the spectra of stroma (ii) and the basement membranes of the ducts (iv) than in the spectra of the stromal cells exhibiting high nucleic acid levels (i) and cells surrounded by adipose tissues (v).

Region II: The $\nu_{as}(CH_2)$ and $\nu_s(CH_2)$ bands at ~2925 and ~2853 cm^{-1}, respectively, are due predominantly to lipids within the cell membranes [31]. The intensities of the $\nu(CH_2)$ bands are significantly higher in the spectra of the basement membrane and adipose tissue (Figures 5.19C (iv) and (v)) than in the spectra of the stromal tissues in Figures 5.19C (i) and (ii) [31].

Region III: The amide I band exhibits high sensitivity to changes in the conformation of the secondary structure of a protein [9]. The amide I and II bands in the spectra of the stromal cells, which also exhibit high levels of bands due to nucleic acids (Figure 5.20(i)) are found at 1655 and 1560 cm^{-1}, respectively. Notably, the positions of these bands were observed at lower wavenumbers in the spectra of the other tissue components. This indicated that the predominant protein secondary structure in the DCIS lesion is that of an α-helix and the other tissues mainly contain proteins, which exist in a β-sheet conformation, as discussed in Section 5.1.3 [9].

The $\nu_{as}(PO_2^-)$ and $\nu_s(PO_2^-)$ bands at 1240 and 1080 cm^{-1}, respectively, are of high relative intensity in the spectra obtained from the basement membrane of the duct, which is composed of epithelial cells, while these bands are of medium relative intensity in the spectra from the stroma and of weak intensity in those from adipose tissues (no cells) [7,15]. No bands were observed in the spectra from the healthy duct lumen area (Figure 5.19C (iii)), and this is consistent with the expected absence of cells.

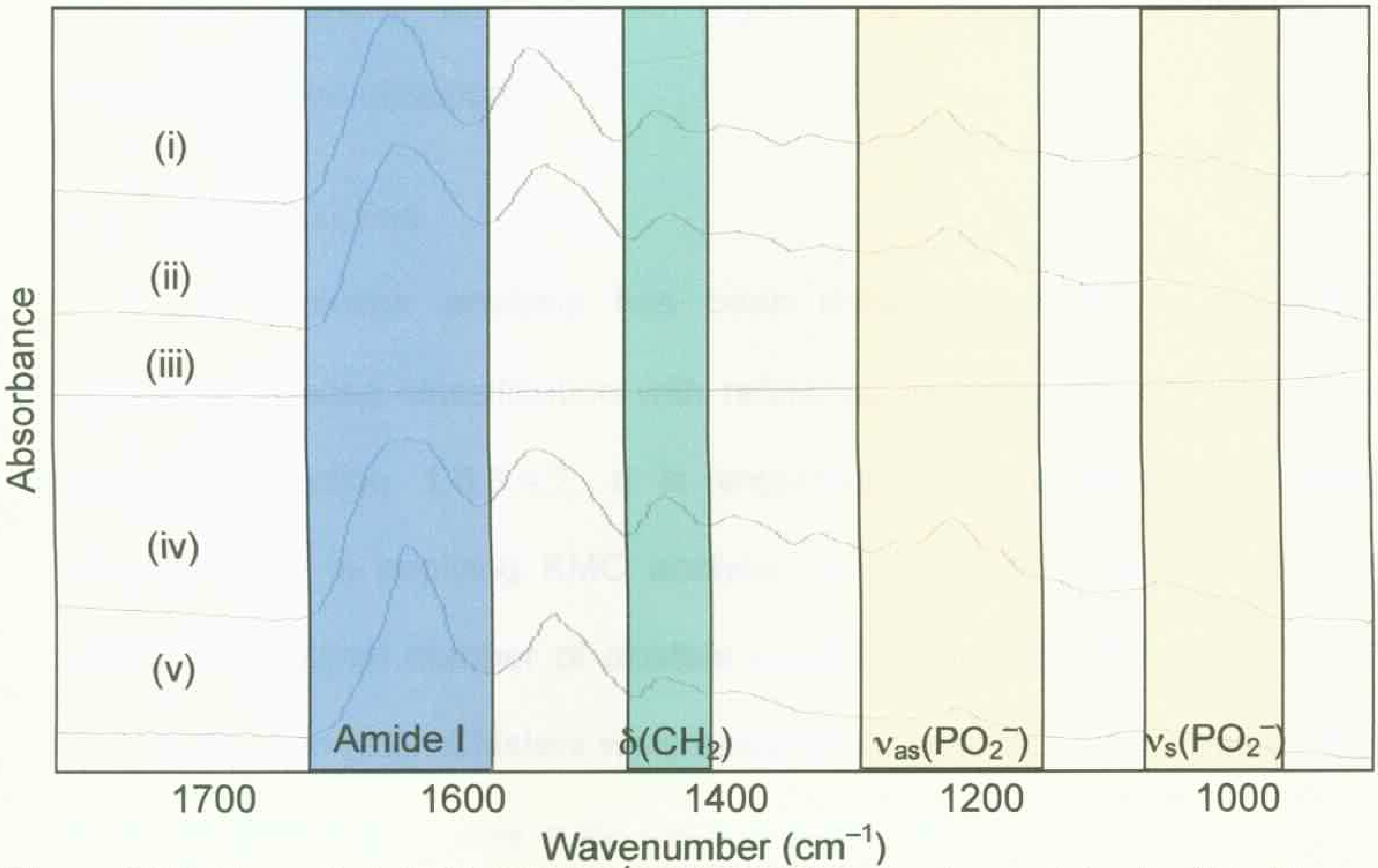


Figure 5.20. Expanded 1800–800 cm^{-1} region of the spectra contained within Figure 5.18 from the normal breast duct lumen and surrounding tissue.

5.3.1.1 NON-HIERARCHICAL CLUSTER ANALYSIS OF NORMAL BREAST DUCT

The functional group map, plotted using data from the 1300–1000 cm^{-1} spectral region, provides information on the distribution and relative concentration of nucleic acids. However, it does not effectively discriminate the morphological features within the tissue, including the distribution of normal and tumour cells [3]. In order to produce maps that could be used for direct comparison with histopathology (adjacent H&E sections), non-hierarchical cluster (NHC) analysis techniques, including K-means cluster (KMC) and fuzzy c-means (FCM) cluster analysis, were implemented. These cluster analyses were applied to the normal breast duct IR image to characterise various tissue components. In KMC analysis, various clusters were selected to determine the optimal groups for breast tissues over the spectral region of 1300-1000 cm^{-1} ($\nu(PO_2^-)$ bands) on the basis of tissue architecture, while FCM analysis was applied to an optimal number of clusters from KMC analysis in various spectral regions (both general and specific regions) to evaluate any additional information could be obtained.

5.3.1.2 KMC ANALYSIS

K-means cluster analysis has been demonstrated to be a useful technique for disease classification with reference to histopathology [21]. As discussed in Section 1.6.3.4.2, it is important to understand the tissue architecture prior to applying KMC analysis (similarly with FCM analysis) to determine the optimal number of clusters suitable to the specific tissue types under investigation [21]. Clusters were assigned to specific tissue types on the basis of histopathology results in the methodology that is described herein.

The dominant feature of the normal breast FFPE section (Figure 5.19) was the duct lumen, which was surrounded by the epithelial cells, basement membranes, and stromal and adipose tissues. KMC analysis using data from the nucleic acid spectral region (1300–1000 cm^{-1}) was applied to the imaging data with as many as five clusters (2C, 3C, 4C and 5C) for tissue classification. The cluster maps are presented in Figure 5.21, together with the mean spectra of each individual cluster from the KMC maps.

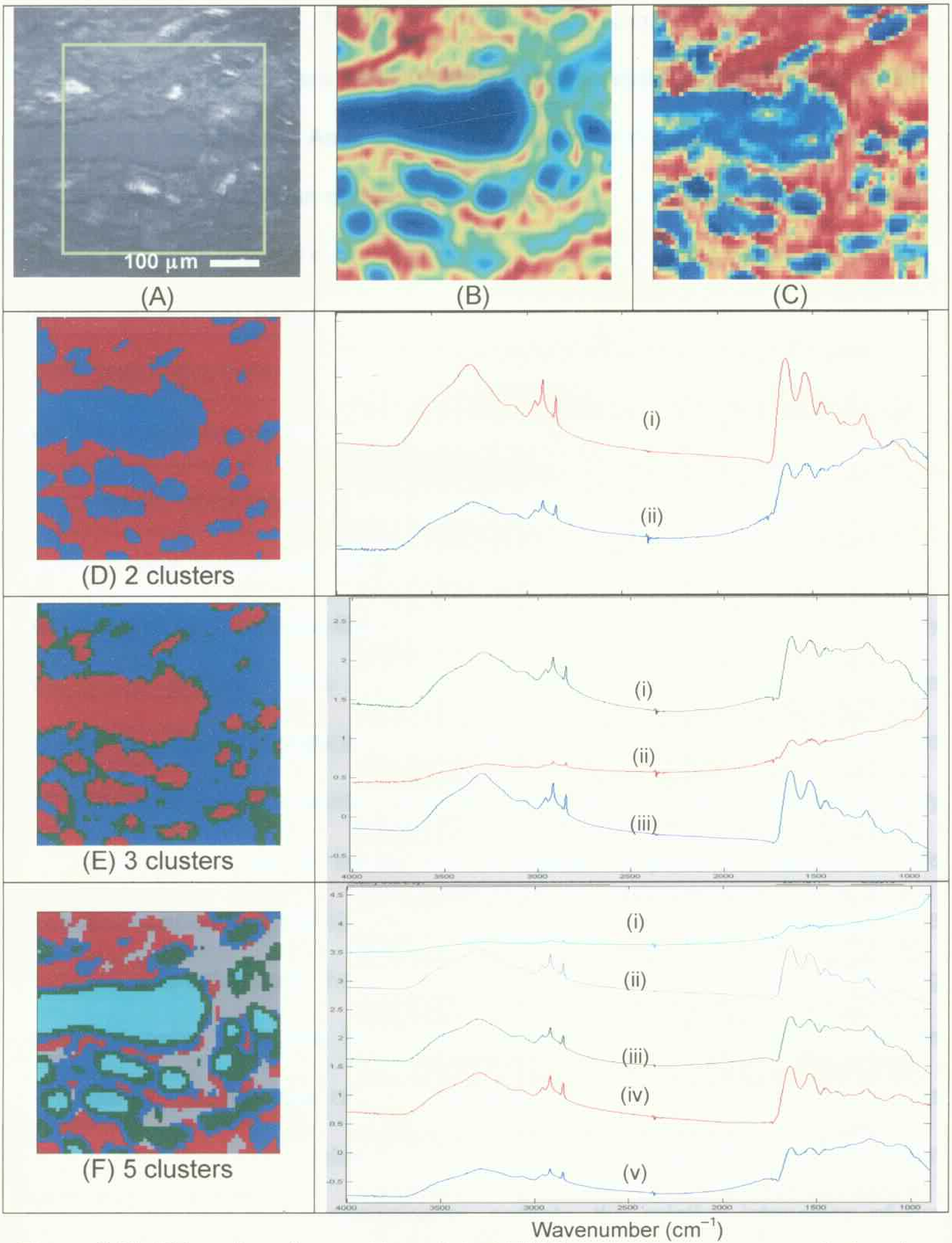


Figure 5.21. IR imaging of normal breast duct FFPE section by k-means analysis using data from the 1300–1000 cm^{-1} region (nucleic acids), together with the mean spectra from each cluster: (A) optical image; (B) area map; (C) vector-normalised area map; (D) 2 clusters; (E) 3 clusters; and (F) 5 clusters. The mean spectra are vector-normalised.

Two clusters clearly distinguish cells (epithelial and stroma, red pixels) and non-cellular components (duct opening and dissolved adipose tissues, blue pixels) in Figure 5.21D, however, the average spectrum of the non-cellular components showed there were biological components (amide bands) present within the blue cluster. As a result, two clusters were not sufficient to classify cellular and extracellular components of the sections, particularly as no cells should be found within the duct openings or within the areas where the adipose tissues (dissolved) existed prior to processing. It is also noted that the blue pixel mean spectrum is affected by an anomalous dispersion artefact [38,53], which is evidenced by the non-linear baseline.

When three clusters were used, the blue pixels in the original 2C KMC analysis were further separated into two additional groups. The duct opening and the dissolved adipose tissues are clustered in red, which are surrounded by the green cluster. The average spectrum of the red cluster contains significantly fewer biological bands, whereas the green cluster spectrum is different from the blue cluster spectrum. This could explain the unexpectedly high intensities of the IR bands from biological components observed in the two-cluster spectrum (blue). When compared with the optical image and the histopathological report, the green region is consistent with areas containing epithelial cells, while the blue cluster is dominated by the stromal tissues (collagens). Comparison of the mean spectra attributed to the epithelial cells (green) and the stroma (blue) clearly shows the spectral differences that exist between different tissue components, notably, the relative intensities of the collagen and nucleic acid bands in the 1300–1000 cm^{-1} spectral region. The red cluster mean spectrum shows relatively fewer biological components were present in the non-cellular regions and was affected by the dispersion artefact.

The KMC cluster map in Figure 5.21F was assembled using five clusters.

All of the spectral clusters were clearly assigned to features present in the histopathological architecture: aqua blue pixels encode the duct openings and dissolved adipose tissues, and green pixels determine the epithelial cells surrounding the openings. Three new clusters (red, dark blue and grey) were separated and were originally from the blue regions of the 3C KMC map. The red pixels are consistent with the areas with high nucleic acid concentration, the grey pixels correlate with areas of stromal tissues, and the dark blue pixels are usually found with the epithelial cell (green) cluster, which correspond to the basement membrane.

In this particular tissue where the duct was the major tissue component, the five-cluster map was sufficient to delineate the major cell types and regions within the tissue. The aqua blue pixel (duct openings and dissolved adipose tissues) mean spectra contained minimal biological peaks, whereas spectral variations were clearly observed in the fingerprint region, especially in the 1300-1000 cm^{-1} region (nucleic acids). With the absence of epithelial cells found within the duct lumen and no apparent cellular invasion through the basement membrane, the duct was classified as normal.

5.3.1.3 FCM CLUSTER ANALYSIS

Fuzzy c-means cluster analysis is another non-hierarchical cluster analytical technique, which partitions objects into groups whose members show a certain degree of similarity [21,54]. Unlike KMC analysis, which displays membership of either 0 or 1, FCM uses a soft linguistic system that has variable membership function values between 0 and 1. The higher-class membership value (1) defines high colour intensity and vice versa [21,54]. Prior to any FCM analysis, it is necessary to understand the histopathological significance of the tissue section under investigation in order to choose the relevant number of clusters for the analysis [3,21]. In addition, the infrared spectrum recorded at each pixel during the experiment provides signature spectra to specific tissue types, e.g., shifts of the amide I band and the intensities of nucleic acid $\nu(PO_2^-)$ bands could differentiate the epithelial cells and the duct openings.

It was determined from the KMC analysis, in the previous section, that five clusters were sufficient to delineate the major cell types of the normal breast duct with reference to the histopathology results. Recent research by Fabian *et al.* have shown that cluster analysis using a wide spectral range (fingerprint region), rather than using specific spectral features, is required for differentiation of different types of breast lesions [42]. The anomalous dispersion artefact observed in the spectral region between 2700–1700 cm^{-1} (Section 4.8) affected the areas where obvious tissue thickness variations occurred, including the edge of the duct openings. However, the mean spectra shown in the KMC analysis did not display this artefact except for a minor effect observed in blue pixel mean spectrum. This observation could be due to the insignificant number of 'edge' spectra recorded in the specific clusters, i.e., most of the spectra were not recorded at the junctions between the duct lumen and the epithelial cells or very thin sample region. Therefore, the use of full spectral region (4000–650 cm^{-1}) in any cluster analysis is not feasible due to the anomalous dispersion artefact, which could prevent the differentiation of tissue types in a sample.

Five-cluster FCM analyses were applied to the normal breast section over the spectral regions of: (A) 3700–3000; (B) 3000–2800; (C) 1900–1000 cm^{-1}; (D) amide I (1750–1580 cm^{-1}); and (E) nucleic acids (1300–

1000 cm^{-1}). The 3700–3000 cm^{-1} and fingerprint (1900–1000 cm^{-1}) regions (Figures 5.22A and C) provide information about the distribution of proteins and nucleic acids, whereas the 3000–2800 cm^{-1} spectral region (Figure 5.22B) separates the tissue components based on the variations in band structure due to lipid $\nu(CH_2)$ modes. Specific FCM images were also plotted for proteins (1750–1580 cm^{-1}, Figure 5.22D) and nucleic acids (1300–1000 cm^{-1}, Figure 5.22E) to differentiate the pathological features of the breast tissue components and thus locate any tumour-affected areas, when referenced to the histopathological results.

The general spectral regions were initially compared to look for trends, prior to investigation of the specific cluster assignments and specific band regions (amide I and nucleic acids). The cluster maps obtained from the data in the 3700–3000 cm^{-1} and the 3000–2800 cm^{-1} regions (Figures 5.22A and B) exhibited spectral similarities in all five clusters, while there were differences observed in the fingerprint region maps, in particular the clusters 1 and 5 (Figure 5.22C). The differences in these clusters may be due to the higher sensitivities to protein and nucleic acid peak positions in the fingerprint region than in the 3800–2800 cm^{-1} region (which may be affected by strong paraffin residual absorptions). In addition, most of the cancer research of the type described in this thesis uses the fingerprint region to discriminate amongst various cellular component and disease states [3,21,26,42].

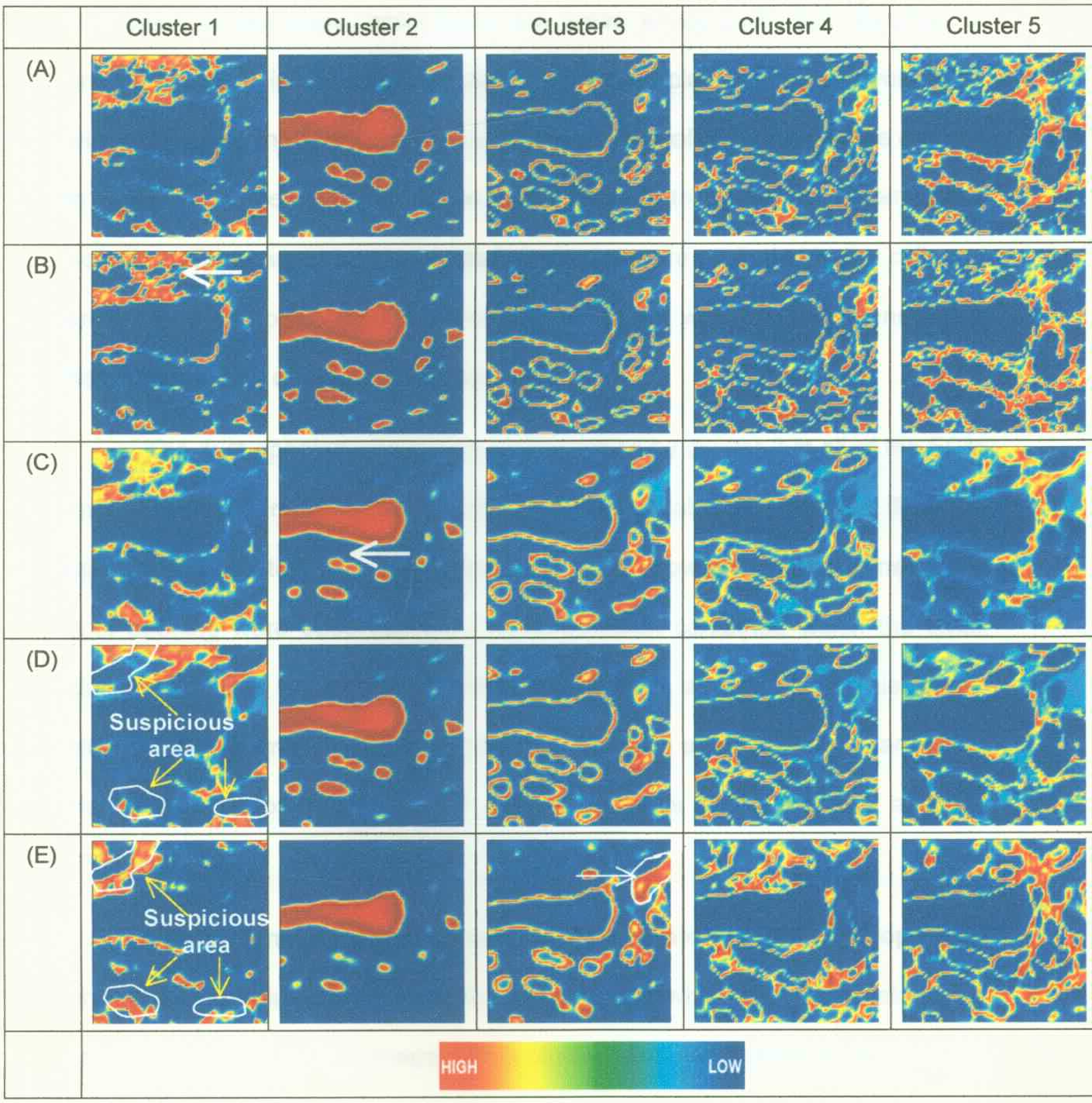


Figure 5.22. Five-cluster FCM cluster images of a normal breast duct plotted from data in the spectral regions: (A) 3700–3000 cm^{-1}; (B) 3000–2800 cm^{-1}; (C) 1900–1000 cm^{-1}; (D) 1750–1580 cm^{-1}; and (E) 1300–1000 cm^{-1}. Experimental conditions are as described in Figure 5.19.

Cluster 1 from the general spectral region FCM maps (Figures 5.22A–C) displayed an area containing a mixture of biological components including proteins, lipids and nucleic acids. When compared with the amide I and nucleic acid maps, there is an obvious concentration of protein and nucleic acids in the top region of the maps, while there are areas where both high densities of protein and nucleic acids were located. When the cluster map was compared with the H&E stained section and the histopathological report, the areas marked in cluster 1 were believed to be suspicious in terms of likely cancerous activities. Other areas on the amide I cluster map are attributed to the collagen protein of the stromal tissues, which are consistent with the deductions made from examination of the optical image.

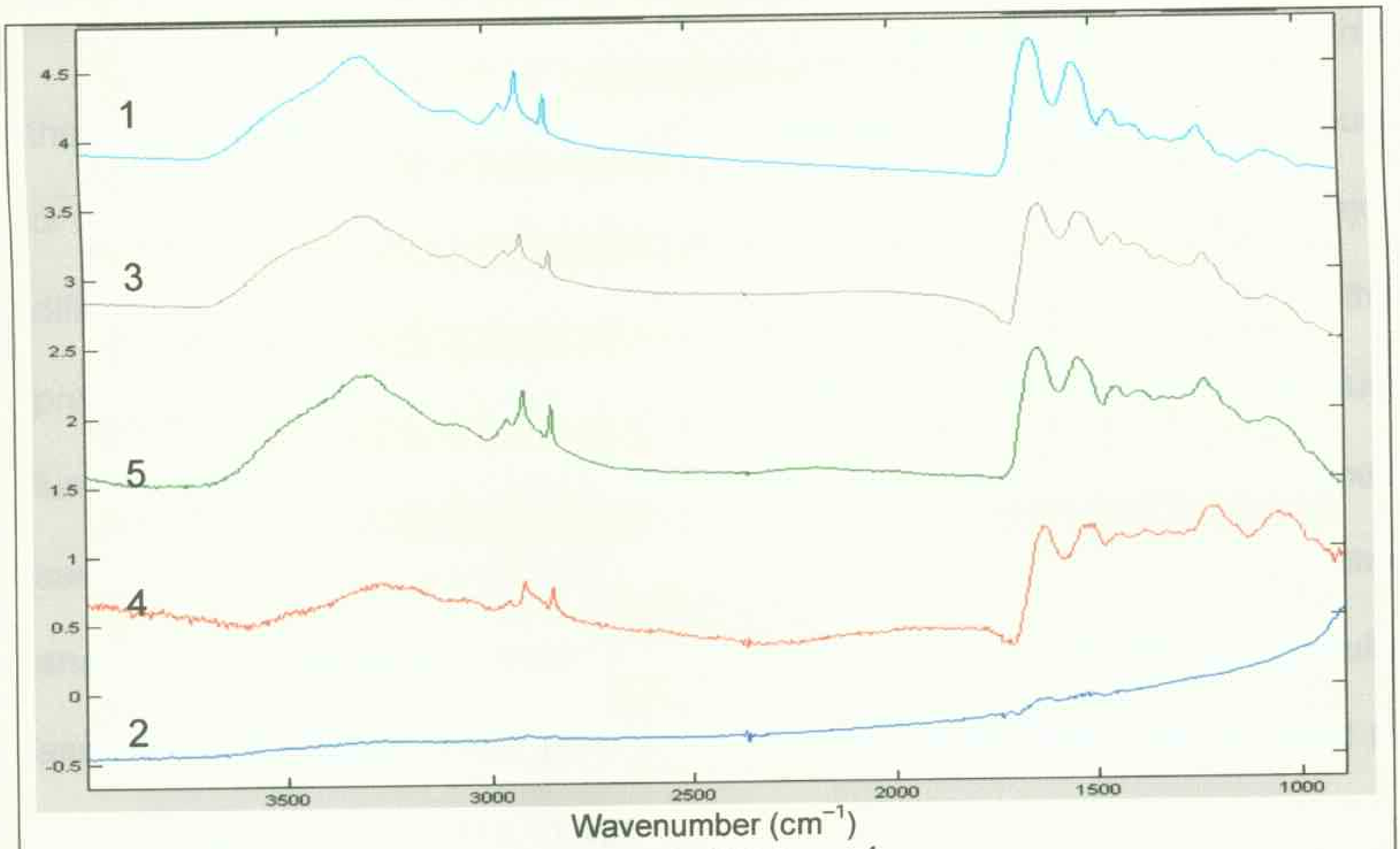
The duct opening was clearly distinguished by cluster 2 in all of the FCM images. Any abnormal cell growth, necrosis or calcification within the duct opening is indicative of a breast condition [52]. However, no cells inside the duct opening were observed from the information contained within the other clusters, which is indicative of a normal breast duct, as expected. This cluster also includes a small contribution from the dissolved adipose tissues (arrows in cluster 2) and correlates with the aqua blue pixels in the 5C-KMC analysis (Figure 5.21F).

Cluster 3 mainly encloses areas that correlate with the duct openings and dissolved adipose tissues, which is consistent with the green pixels in the 5C-KMC and the optical image. This cluster also has contributions from the epithelial cells (both inner cuboidal epithelial and outer myoepithelial cells) surrounding the duct lumen. Although a collection of epithelial cells is located outside of the duct opening, they are not clustered with the suspicious cells (cluster 1). In fact, the pathology report confirms that these cells are non-cancerous. The area marked in Figure 5.22 shows high epithelial cell activity but a low-to-medium intensity in a smaller area (arrow) is observed in the amide I image (Figure 5.22D). This may be due to the presence of collagens, which also have bands in the same spectral region (1300–1000 cm^{-1}) [26,31].

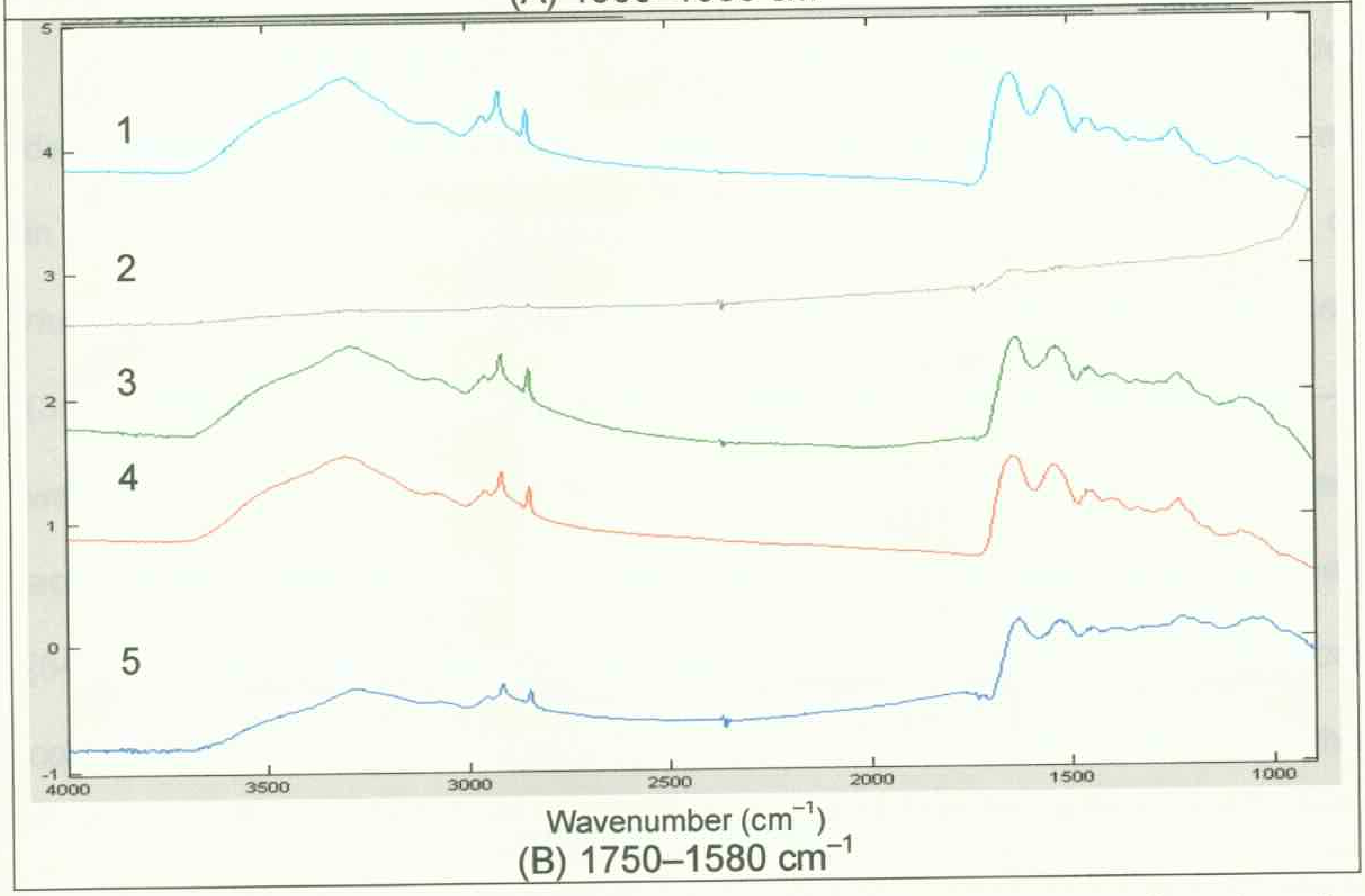
The basement membrane of the duct is grouped as cluster 4; it envelops the epithelial cells of the duct in cluster 3. The intra- and inter-lobular stromal tissues are the major breast supporting components, and are grouped together in cluster 5. The stromal tissues are evenly distributed across the section and their IR bands are found in the areas that do not contain any epithelial cells.

The mean spectra from five-cluster FCM analysis using the spectral regions of 1900–1000 cm^{-1} and 1750–1580 cm^{-1} are presented in Figure 5.23. The duct opening (cluster 2, blue) mean spectrum contains no biological signature peaks (protein, nucleic acids and lipid). Meanwhile, spectral variations are observed in different tissue components within the fingerprint region in particular the $\nu(PO_2^-)$ and amide bands. It is also noted that a subtle anomalous dispersion artefact was observed in the spectra from clusters 3 (grey) and 4 (red) in Figure 5.23A, while only a minor contribution from this dispersion artefact was observed in cluster 5 mean spectrum (blue) in the 1750–1580 cm^{-1} spectral region in Figure 5.23B.

The normal breast duct images plotted using the functional group areas, k-means clustering, and fuzzy c-means clustering were compared. The functional group maps illustrated the distribution of biomolecules in the tissue lesion but was unable to determine the specific tissue components. These components, including duct lumen, epithelial cells, basement membranes, adipose tissues (dissolved) and connective tissues, were successfully classified using KMC and FCM cluster maps. Five clusters were shown to represent the majority tissue components and of these two clustering techniques, KMC produced maps that are simpler to interpret and well-correlated to the histopathology results.



(A) 1900–1000 cm^{-1}



(B) 1750–1580 cm^{-1}

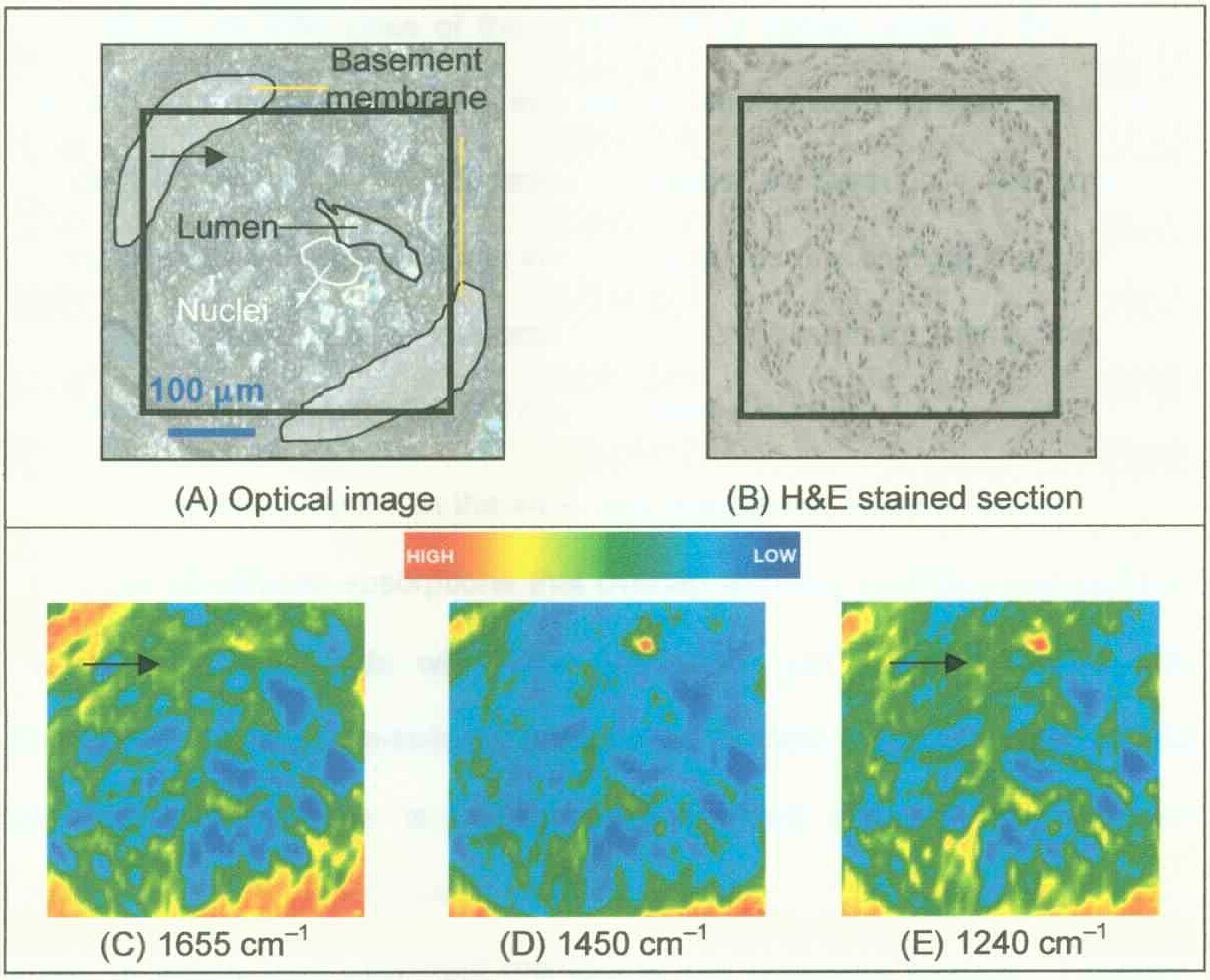
Figure 5.23. Mean spectra extracted from five-cluster FCM analyses based on the spectral regions of: (A) 1900–1000 cm^{-1}; and (B) 1750–1580 cm^{-1}. Numbers correspond to the specific clusters extracted in the spectral region.

5.3.2 IR IMAGING OF A CRIBRIFORM DUCTAL CARCINOMA *IN-SITU*

Ductal carcinoma *in-situ* (DCIS) is a pre-invasive cancer localised within the breast ducts [55]. As epithelial cells proliferate, the architecture of the duct or the terminal ductal-lobular unit (TDLU) of the breast alters [56,57]. The major difference between a low-grade DCIS and a normal or benign duct is the presence of tumour cells arranged in a cribriform architecture within the duct lumen [56,57]. Fabian *et al.* examined two high-grade DCIS cryosectioned samples using FPA-FTIR imaging in combination with hierarchical cluster analysis. It was demonstrated that the tumour cells could be successfully assigned to the same colour pixel within the duct lumen, which corroborated to the H&E stained section [10,11,42].

The photomicrographs of an unstained cribriform DCIS (low-grade) deparaffinised FFPE section and an adjacent H&E stained section investigated in this study are presented in Figures 5.24A and B, respectively. A selection of nuclei and duct openings are highlighted and the box indicates the area imaged (350 × 350 μm^2). The functional group maps presented in Figures 5.24C–E were plotted by calculating the areas of the protein (amide I), lipid and nucleic acid bands centred at 1655, 1450, and 1240 cm^{-1}, respectively, using Cytospec [54]. These maps illustrate the relative distribution of these biological components. Thus, the functional group maps can be used together with the unstained and H&E stained sections, to identify regions with a high density of protein and nucleic acids that can be correlated to regions of high cell proliferation. A normal breast duct should not contain any cells; any spectral evidence of biological components within the duct is, therefore, indicative of abnormalities.

The protein (amide I) functional group map (Figure 5.24C) shows that a network-shaped feature of medium intensity (green) is found in the centre of the duct, and is surrounded by another feature displaying a medium-to-high intensity (red) on the peripheral region. When compared with the H&E stained section, the peripheral region is attributed to the basement membrane, while the network-shape feature is due to the tumour cells within the duct lumen arranged in a DCIS cribriform architecture.



(A) Optical image (B) H&E stained section

(C) 1655 cm^{-1} (D) 1450 cm^{-1} (E) 1240 cm^{-1}

Figure 5.24. Images of a 5-μm thick deparaffinized FFPE DCIS section. Optical images of: (A) unstained; and (B) adjacent H&E stained sections. The area imaged using FTIR spectroscopy (C–E) is marked by the black square. IR area maps centred at: (C) amide I (1655 cm^{-1}); lipid (1450 cm^{-1}); and (E) nucleic acids (1240 cm^{-1}). Experimental conditions: 64 × 64 array (4096 spectra); 256 scans; 4 cm^{-1}; ×15 magnification; image area: 350 × 350 μm^2.

Lipids are usually present in the breast in the form of adipose tissue and within the basement membrane [8,29]. The presence of lipids in the duct (green) as shown in Figure 5.24D is due to the lipoprotein from the tumour cell membranes. The medium-to-high intensity of lipids on the periphery correlates with the protein map, shown in Figure 5.24C, which is attributed to the basement membrane. Confirmation that this region is that of basement membrane is made on comparison of the image with the H&E stained section.

The unstained and H&E stained cribriform DCIS sections clearly show there are tumour cells contained within the duct. The presence of nuclei was deduced by the intensities of the bands due to nucleic acids in the IR map. Figure 5.24D reveals a medium intensity of bands due to nucleic acids within the duct lumen, which are indicative of cellular proliferation. The periphery, again, displays a medium-to-high intensity of bands due to nucleic acids, which have the histopathological characteristics of the basement membrane of the duct containing myoepithelial and cuboidal epithelial cells.

Collagen is present in the intra- and interlobular stromal tissues and has a number of infrared absorptions that overlap with the $\nu_{as}(PO_2^-)$ and $\nu_s(PO_2^-)$ modes of nucleic acids within the 1300–1000 cm^{-1} spectral region [58]. Therefore, the medium-to-high intensity of nucleic acids, located on the periphery of the image, is likely to be associated with both the basement membrane and stromal tissues. Similarly, the high concentration also indicates areas of stroma and basement membrane that physically overlap. However, the difference between stroma and basement membranes is not obvious from the functional group maps.

A normal breast duct is composed of a bi-layer of myoepithelial and cuboidal epithelial cells that are surrounded by basement membranes [33]. The ducts and lobules are supported by the intra- and inter-lobular stromal tissues

[33]. As indicated by arrows in Figures 5.24C and E, the distribution of the protein and nucleic acids indicated a sieve-like architecture of cellular components present within the duct lumen, which is highly correlated with the adjacent H&E section (Figure 5.24B) [33]. Although functional group maps provide information on the distribution of bands attributed to proteins, nucleic acids and lipids over the mid-infrared region, direct tissue classification cannot be undertaken. In order to improve the diagnostic outcome, KMC and FCM analyses were undertaken, the results of which are discussed in detail below.

5.3.2.1 KMC ANALYSIS

In order to identify the morphological features of the DCIS breast lesion, the amide I band area map (Figure 5.25A) together with three-cluster, five-cluster and six-cluster KMC images based on the amide I region (Figures 5.25B–D) were produced. The mean spectra extracted from each cluster analysis are also presented.

The components in the three-cluster KMC image (Figure 5.25B), were assigned by cross-referencing the clusters to the H&E stained section as follows: (i) the basement membranes (green); (ii) duct openings (red); and (iii) epithelial tissues (blue). The basement membrane cluster clearly correlates with the medium-to-high intensity of the protein amide I image (Figure 5.25C), as do the well-defined epithelial tissues within the breast duct. The red and blue pixels strongly correlated with the unstained section (Figure 5.25A) where epithelial cell nuclei (blue) are located within the duct lumen (red) and surrounded by basement membranes (green cluster).

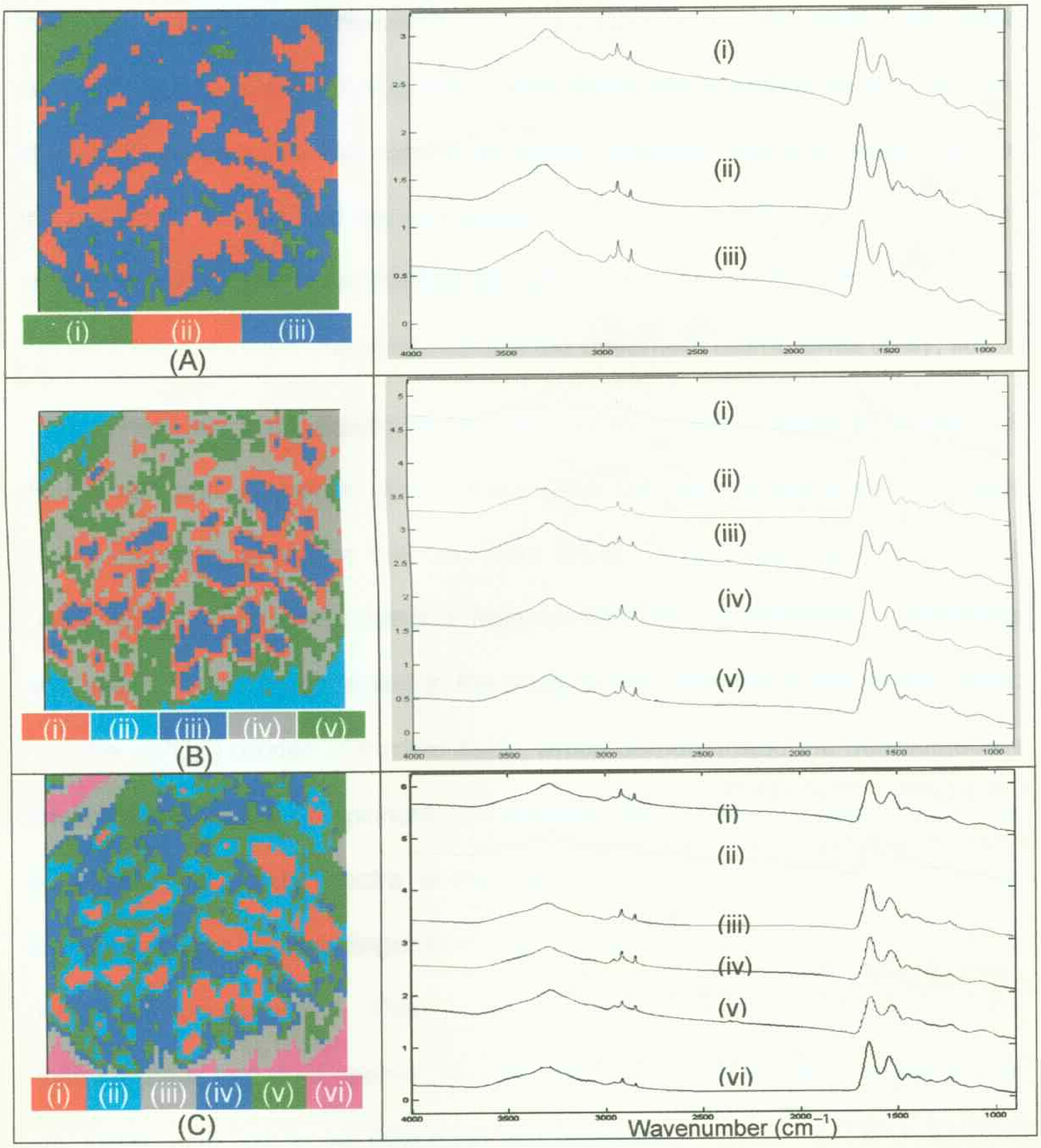


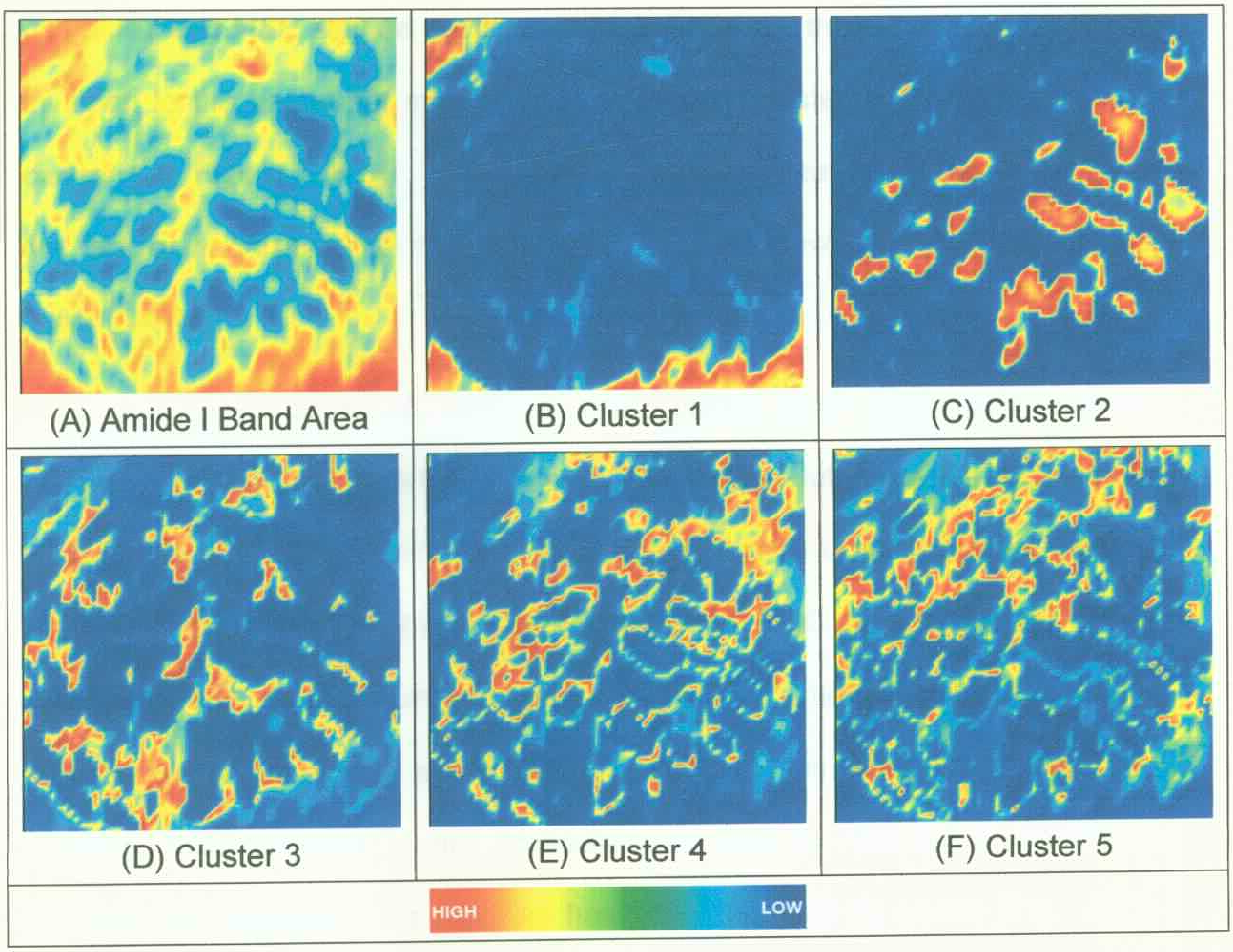
Figure 5.25. KMC maps of the vector-normalised amide I band area and the mean cluster spectra of the deparaffinised FFPE cribriform DCIS section using: (A) three clusters; (B) five clusters; and (C) six clusters.

A five-cluster KMC analysis (Figure 5.25C) was applied which discriminated the epithelial tissue pixels from 3-C KMC image (blue, Figure 5.25A) into three clusters. Two of these three clusters represent the epithelial tissues: the green cluster (v) represents the myoepithelial tissues attached to the basement membranes and the red cluster (i) represents the cuboidal epithelial tissues enveloping the duct openings. The grey cluster (iv), the last of these three clusters, corresponds to the intralobular stroma, which are loose connective tissues surrounding the lobules within the cribriform duct. The blue cluster (iii) defines the duct openings, which correlate with the morphology of the unstained section and the red cluster (i) in Figure 5.25D. Similarly, the six-cluster image further differentiated the basement membranes into two groups: the connective tissues (magenta, vi) and the basement membranes (grey, iii).

There were no significant changes in the clusters assigned within the duct lumen and, therefore, five clusters were sufficient to delineate the major tissue components within the cribriform DCIS. The mean spectra from all cluster maps have similar spectral features attributed to biological components with subtle variations observed in the position and intensity of the amide I band and the $\nu(PO_2^-)$ modes of nucleic acids, which demonstrated the non-subjective advantage of tissue component classification using KMC. Finally, it must be noted that the mean spectra of the duct opening clusters in all KMC maps exhibited peaks in the fingerprint region that are due to the presence of biological components. This mean IR spectrum contained a mixture of epithelial cells and duct openings, which was due to the limitation of the spatial resolution (< 15 μm) of the FPA FTIR instrumentation.

5.3.2.2 FCM ANALYSIS

Five-cluster FCM images based on the amide I band area (1750–1550 cm^{-1}, Figure 5.26) were plotted and are presented in Figures 5.26B–F. As previously described, five clusters best illustrated the distribution of the individual tissue components within the DCIS duct and included: the basement membrane, epithelial cells (myoepithelial and cuboidal), intralobular stroma and also any content presence within the duct openings (calcification or lactation).



(A) Amide I Band Area

(B) Cluster 1

(C) Cluster 2

(D) Cluster 3

(E) Cluster 4

(F) Cluster 5

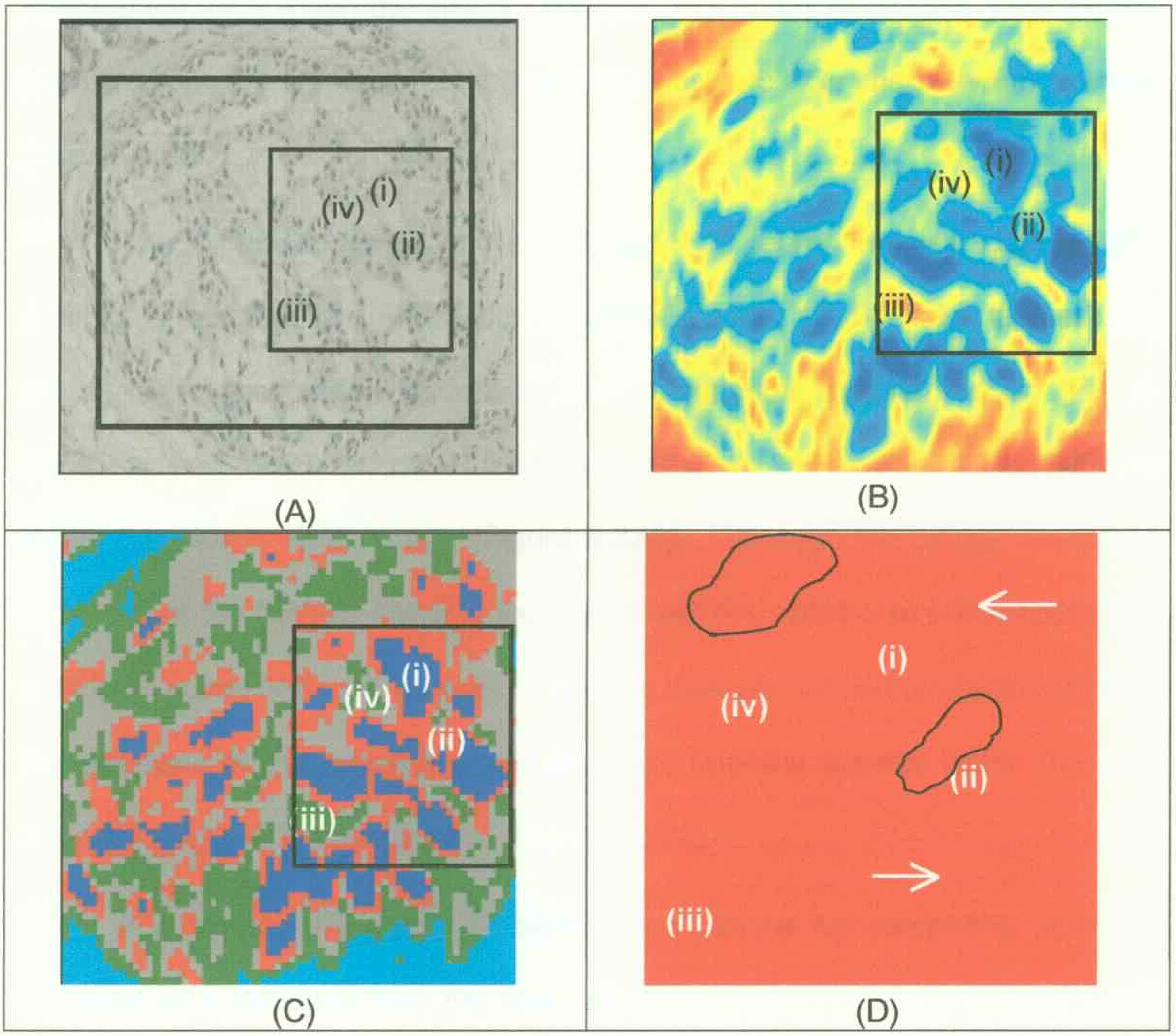
Figure 5.26. (A) Amide I band area functional group map and FCM analysis IR images of the cribriform DCIS duct based on the amide I band areas using five clusters: (B) cluster 1: basement membranes; (C) cluster 2: duct openings, (D) cluster 3: high concentration of tumour cells; (E) cluster 4: tumour cells with intralobular stromal tissues; and (E) cluster 5: intralobular stroma supporting the tumour cells.

High-grade comedo DCIS lesions are associated with the breakdown of the myoepithelial cell layer and basement membrane, and the presence of necrosis (the centre of the duct is plugged with dead cells) [59].. This was previously investigated by Fabian *et al.* who successfully identified necrosis within the duct lumen but no apparent invasion through the basement membrane [42]. There are three different tissue classes within the cribriform DCIS, two of which were assigned to the myoepithelial (Figure 5.26D) and cuboidal epithelial (Figure 5.26E) cells, whereas the intralobular stromal tissues grouped in their own cluster (Figure 5.26F). The duct openings clustered in one group as shown in Figure 5.26C, which correlates to the area marked on the unstained and H&E stained sections. In this instance, there is no apparent cell invasion to the basement membranes (aqua blue cluster in Figure 5.25B and cluster 1 in Figure 5.26B). The absence of comedo features of the TDLU and the lack of necrosis within the duct lumen are indicative of a low-grade DCIS lesion [44-48].

In comparison to KMC analysis, interpretation of the FCM cluster maps is not straightforward as reported by Lasch, since it is a soft linguistic approach in which there is a internal membership variation (0 to 1) in each cluster, instead of the hard class membership of either 0 or 1 in the KMC analysis [21,54]. If the number of clusters was further increased (>5), the interpretation of the FCM maps would become more ambiguous in terms of known tissue composition. In addition, FCM requires longer computational time for the same conditions (clusters and regions) than KMC as has been reported by Lasch [21]. As a result, KMC analysis is the preferable analytical tool for preliminary spectral classification prior to developing a complex artificial neural network.

5.3.2.3 PCA DISPLAYING THE TUMOUR NUCLEI WITHIN THE DCIS DUCT

Principal component analysis is another statistical analysis technique provided by Cytospec. It differentiates spectral variations by reducing the data model to the least number of PCs [60]. A PCA image based on the amide I band area was applied to the DCIS section to model the data with the least number of PCs and the first three PCs represented the majority of the total variance (> 98%). Figure 5.27 displays: (A) the H&E stained section and three amide I IR images of the same DCIS duct that are plotted on the basis of: (B) the area underneath the amide I band at 1655 cm^{-1}; (C) a five-cluster KMC image; and (D) the vector-normalised second-derivative PCA of the area marked in the DCIS section.



(A) (B)

(C) (D)

Figure 5.27. Comparison of various IR image analyses plotted against the amide I band: (A) H&E stained section; (B) band area 1750–1550 cm^{-1}; (C) five-cluster KMC map; and (D) enlargement of the image of interest in (A) using vector-normalised second-derivative principal component analysis (overlaid PC2 and PC3).

The distribution of the protein (based on the area of amide I band, Figure 5.27B) and the five-cluster KMC image (Figure 5.27C) in the Section 5.3.2.1 were correlated to the histopathological H&E stained section (Figure 5.27A). The sieve-like structure within the duct is due to the tumour cells but there is no evidence to suggest invading cells at the basement membrane. PCA is a powerful technique to differentiate IR spectra of breast lesions, as discussed in Section 5.2.4.1. However, it is not a commonly used technique for analysing FTIR images compared to KMC, FCM and hierarchical cluster analysis, which is probably due to the limited functions available in Cytospec [54]. The PCA of the spectral variance within the amide I spectral region was examined, then the PC scores (PC2 *vs.* PC1) were plotted, and the best two PCs images were overlaid (Figure 5.27D).

The expanded region of the PC2 and PC3 scores image (Figure 5.27D) exhibits areas that are attributed to epithelial cells, non-cellular tissues and voids. This observation is consistent with the H&E stained section (Figure 5.27A) as well as the red pixels (epithelial cells) and blue pixels (duct openings) in the five-cluster KMC image (Figure 5.27C). Although the spatial resolution (~15 μm) of the FPA-IR imaging technique was not capable of differentiation of individual nuclei (< 5 μm), regions containing dark red pixels compared well with the positions of the cribriform epithelial cells (arrows) present within the duct lumen in the H&E stained section (Figure 5.27A).

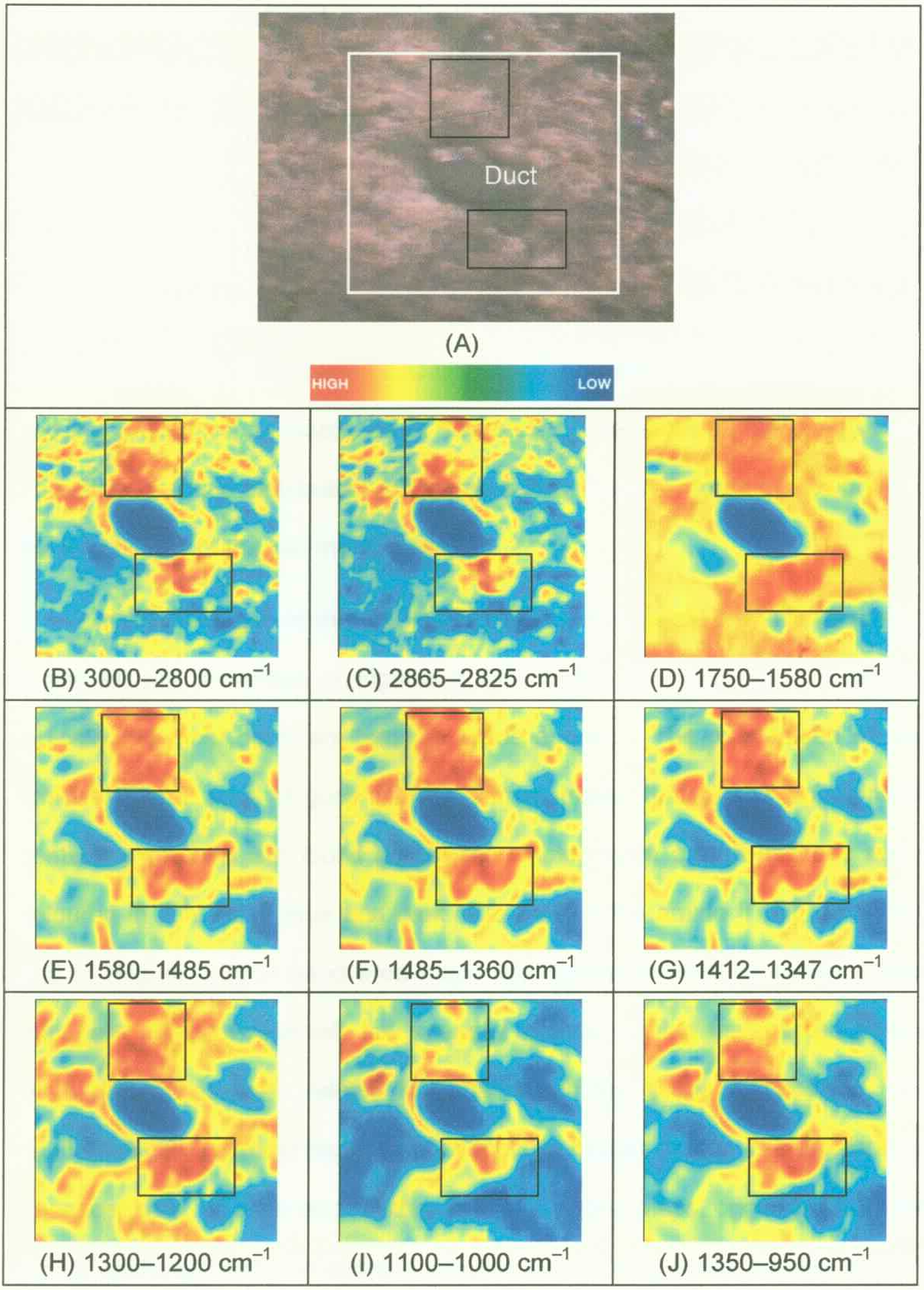
It is particularly useful to classify the lesions by comparing both the morphological features from the histopathology section of the cribriform DCIS and the spectral distribution of various cellular components in the IR images. Different areas (i to iv) shown in the histopathology section (Figure 5.27D) were related to the KMC image in Figure 5.25C and the optical image of the unstained section (Figure 5.24A). Multiple layers of epithelial cells (marked) were observed in the PCA image; these aggregations are indicative of abnormalities.

By combining the functional group map with the multivariate analytical images (KMC, FCM and PCA), both spectral and pathological information are obtained and thus this combination of techniques can be used in parallel with histopathology to increase the diagnostic accuracy for detection of DCIS lesions, as well as to ascertain whether they remained localised within the duct. The low-grade cribriform DCIS examined in this research using KMC, FCM and PCA and the high-grade DCIS reported by Fabian *et al.* clearly demonstrate that a combination of FPA-FTIR imaging and MSA could locate tumour cells within a DCIS lesion [42].

It is deduced that for routine tissue classification, KMC analysis using between three and five clusters, over the fingerprint region, provides the necessary diagnostic information for tissue classification with high sensitivities when referenced to the histopathology results. The functional group area or intensity maps can be used to illustrate the distribution of the biological components across the tissue but cannot be used for classification. Meanwhile, the interpretation of FCM cluster maps was typically complicated due to the soft-linguistic approach as mentioned previously by Lasch and co-workers [21]. Similarly, excessive numbers of clusters used in KMC and FCM analyses on the imaging data are not practical as there are only limited tissue components in any given tissue section [21].

5.3.3 INVASIVE DUCTAL CARCINOMA

Tumour cells are localised within the ducts in DCIS whereas those from invasive ductal carcinoma (IDC) invade beyond the basement membranes of the duct to the surrounding tissues [52]. Therefore, high cellular activity would be expected to be observed in the basement membranes of a malignant lesion. IR functional group maps of an IDC lesion were constructed using the diagnostic band areas of: (i) the $\nu(CH)$ region (Figures 5.28B and C); (ii) amides I, II and III (Figures 5.27D, E and G); (iii) lipid $\delta(CH_2)$ (Figure 5.27F); and (iv) the nucleic acid $\nu(PO_2^-)$ (Figures 5.28H-J).



(B) 3000–2800 cm^{-1} (C) 2865–2825 cm^{-1} (D) 1750–1580 cm^{-1}

(E) 1580–1485 cm^{-1} (F) 1485–1360 cm^{-1} (G) 1412–1347 cm^{-1}

(H) 1300–1200 cm^{-1} (I) 1100–1000 cm^{-1} (J) 1350–950 cm^{-1}

Figure 5.28. (A) Photomicrograph of an unstained 5-μm-thick IDC breast FFPE section and (B–J) IR images plotted using band areas at different spectral regions. Experimental conditions: 512 scans; 4 cm^{-1} res.; ×15 magnification; map dimension: 350 ×350 μm^2, 4096 total spectra.

When the functional group maps were compared with the optical image (Figure 5.28A), the low intensities (blue) found in all functional group maps correspond to the duct opening. All functional group maps displayed relatively high intensities in the regions marked in black boxes could be assigned to the areas comprised of epithelial tissues. Lipids are the major component of adipose tissues (long-chain fatty acids) and cell membranes (phospholipid and lipoprotein) [8,29]. The imaged IDC lesion was prepared using FFPE sectioning and this procedure, as discussed in Section 2.2.3, removes all adipose tissues. Therefore, the lipid distribution in Figures 5.28B, C and F could be assigned to the lipid bilayer of the cell membranes.

Infrared images of the IDC lesion based on the protein bands (amides I, II and III) are presented in Figures 5.28D, E and G. The amide I band is typically the most dominant band in tissue spectra. The map based on the amide I band area (Figure 5.28D) demonstrated that tissue components containing proteins are found across the whole region except in those areas coloured in light aqua blue and dark blue. When the amide II (Figure 5.28E) and III (Figure 5.28G) band area maps are compared with the amide I map, there are areas of relatively low intensity located outside the marked boxes. Such differences were related to variations in the tissue composition, i.e., cancerous cells (marked boxes) and connective tissues (else). The proliferation of cancer cells, as seen on the other functional group maps, originated from the ductal epithelial cells and invaded into the surrounding breast tissues through the basement membrane.

The images depicting the nucleic acid distribution (Figures 5.28H and I) display elevated intensities in the marked regions, which are due to the cell proliferation beyond the basement membrane of the duct into the surrounding stromal tissues, as expected. Figure 5.28J, which displays maps of the 1350–950 cm^{-1} diagnostic area, including bands due to nucleic acids and carbohydrates (1200–1000 cm^{-1}), also exhibits a similar cellular distribution as is seen in the other functional group maps. Therefore, the marked areas are likely to be due to cell proliferation originated from the IDC duct. These results correlate with the classification made by the pathologist after examination of the adjacent H&E stained section.

5.3.3.1 KMC ANALYSIS

K-means cluster maps using three or five clusters, were plotted based on the vector-normalised 1800–950 cm^{-1} region of the IDC IR image, and the results are depicted in Figures 5.29A and B. Here, the three-cluster KMC map divides the IDC lesion into recognisable regions associated with the optical image and functional group images. The green pixels (region I) correspond to areas with a high concentration of tumour cells. The red pixels (region II) represent regions containing no cells – i.e., the duct openings, and regions with a significantly low intensity of proteins (very thin sample thickness), and the blue pixels (region III) are cells with slightly lower intensities of proteins (amides I and II) and nucleic acids, including normal epithelial and stromal tissues. The mean spectra from both green (I) and blue (III) clusters exhibited bands due to biological components with subtle spectral variations (band shifting and intensities) mainly in the amide I band and in the nucleic acids band region (1300-1000 cm^{-1}). However, peaks attributed to biomolecules (amides and nucleic acid bands) were observed in the mean spectrum of the non-cell cluster (II), which indicates there are cellular components in the parts of the tissues that map onto the red pixels. Therefore, a three-cluster KMC analysis is not sufficient to delineate the major tissue content.

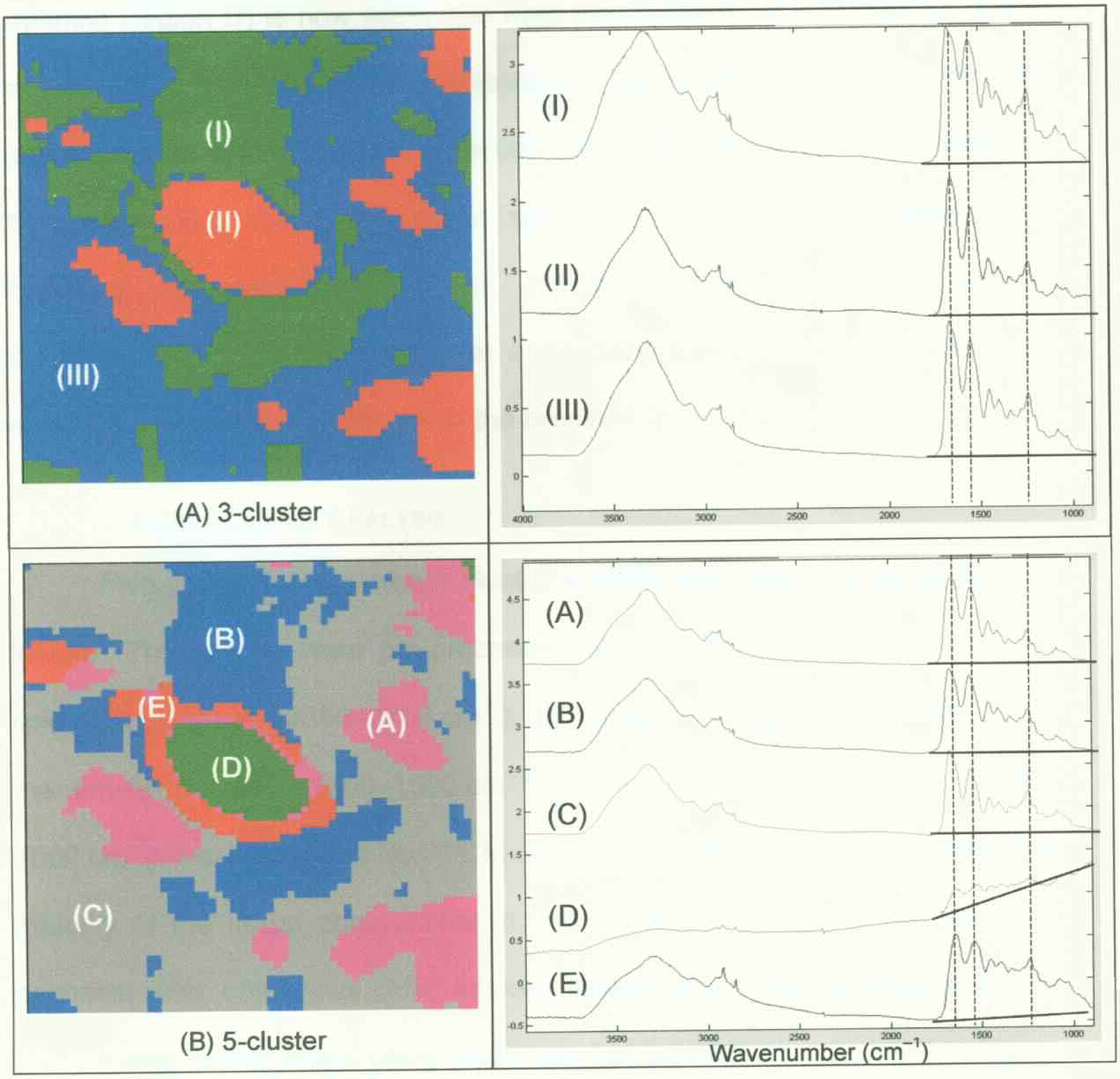


Figure 5.29. K-means cluster maps and the mean cluster spectra from a IDC lesion using: (A) three-and (B) five-cluster methods in the 1800–950 cm^{-1} spectral region. Individual clusters are marked with I, II or III in three-cluster map; and A, B, C, D or E in five-cluster map.

The five-cluster KMC map further divides region I of the three-cluster map (Figure 5.29A), the epithelial cell components (green), into two components (regions E - red and B - blue). Region E is associated with the epithelial cells surrounding the duct lumen, while region B is associated with the proliferated epithelial cells that have invaded the surrounding breast tissues. In addition, the red pixels (II) in the three-cluster map are now subdivided into two groups (regions D - green and A - magenta) in the five-cluster map. The duct opening (region D) is now separated from the region A, which could be due to an uneven surface, as there is evidence of bands attributed to biomolecules present in the mean spectrum. Although cell proliferation is seen to invade into the surrounding tissue, the mean spectrum of region D indicates some biological component is still present. This may indicate that tumour cells are contained within the duct lumen at a relatively low concentration and have spread across to other ducts within the breast.

5.3.3.2 FCM CLUSTER ANALYSIS

Five-cluster fuzzy c-means cluster analysis was used to analyse the IDC lesion. The spectra were pre-processed using vector-normalisation, and the lipid $\nu(CH_2)$ region (3050–2800 cm^{-1}), the fingerprint region (1800–950 cm^{-1}), the amide I region (1750–1580 cm^{-1}) and the nucleic acid region (1300–1000 cm^{-1}) were selected for FCM analysis. Five clusters represent the majority of the tissue components in this specific IDC region (duct opening, epithelial cells, cancerous cells, adipose tissues and connective tissues), and the resulting cluster maps using these spectral regions are presented in Figure 5.30. If the number of clusters was further increased (>5), the interpretation of the FCM maps became more ambiguous in terms of the known histopathology results (data not shown).

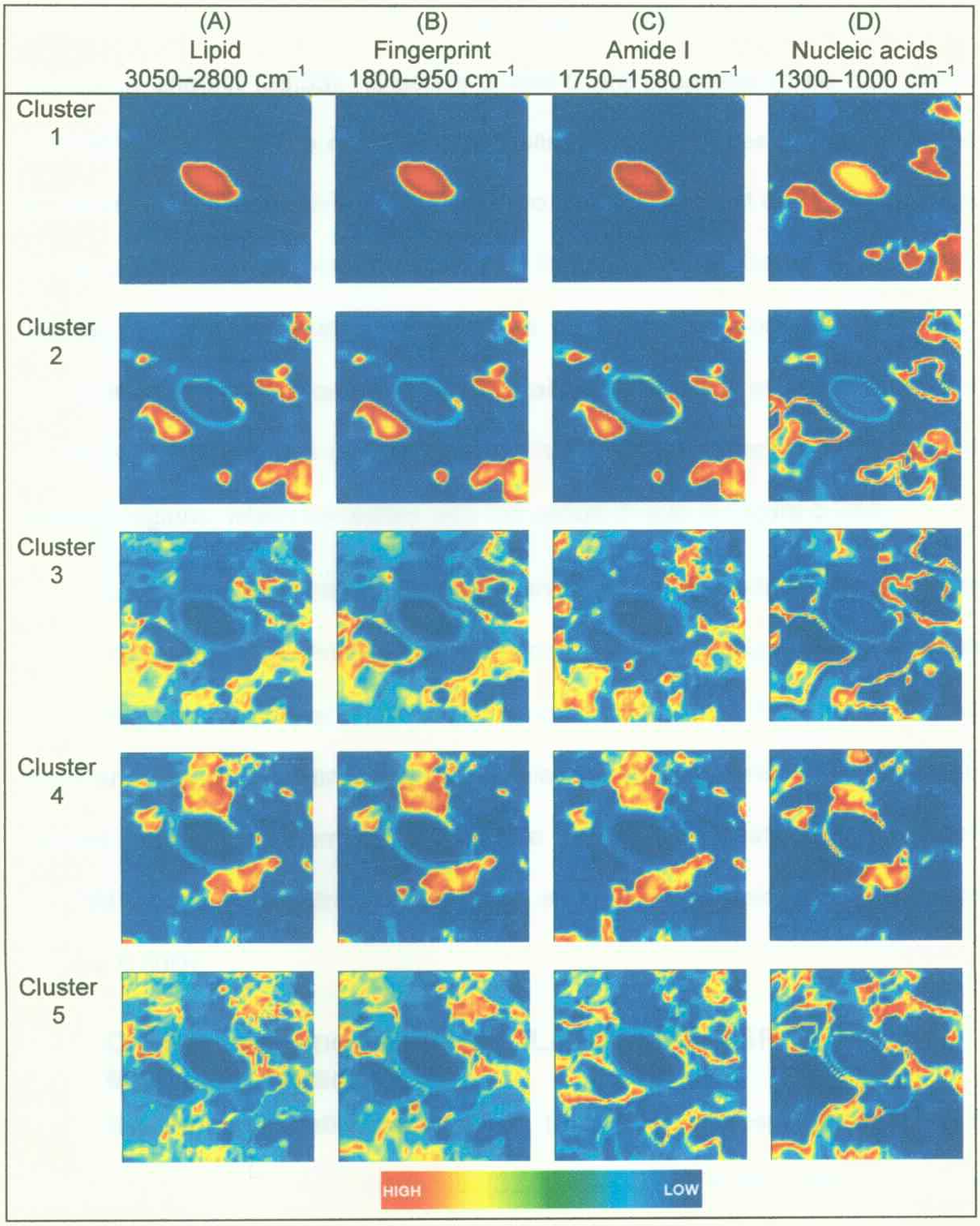


Figure 5.30. Five-cluster FCM IR area images from Figure 5.28 using the vector-normalised regions due to the: (A) lipid ν(CH_2) band; (B) fingerprint bands; (C) amide I band; and (D) nucleic acid bands. Clusters 1 to 5 refer to specific tissue components.

When referenced to the histopathology report, cluster 1 highlights the duct opening in all spectral regions (Figures 5.30A–D). Extra features are observed in the nucleic acid cluster image (Figure 5.30D). These extra features are also observed in cluster 2 of other spectral region maps, and are due to the 'uneven' surfaces of the sections that may be affected by the anomalous dispersion artefact [38,53]. Clusters 3 and 5 are proposed to be due to the stromal tissues, and are dominated by collagen (protein), as is evident in the amide I map. The reason that there are two clusters, and not one, could be due to the subtle spectral variations in the loosely and tightly packed stromal tissues. Regions previously identified as cancerous by comparing the five-cluster KMC image (region D) with H&E stained section are evident in cluster 4 of the FCM image. This cluster clearly exhibits strong intensities in all imaged spectral regions, which correlates with the optical image in Figure 5.28A.

Choosing the correct cluster analysis is important for spectral interpretation and classification of tissue components to ensure high correlation exists between the cluster and histopathology results. Here, KMC images using three and five clusters provided the simplest direct comparison with the H&E stained sections. Furthermore, this sample demonstrates that the interpretation of FCM images is not straightforward, as seen in the nucleic acid FCM maps (Figure 5.30D).

5.4 CHARACTERISATION OF BREAST LESIONS USING SR-FTIR MICROSPECTROSCOPY

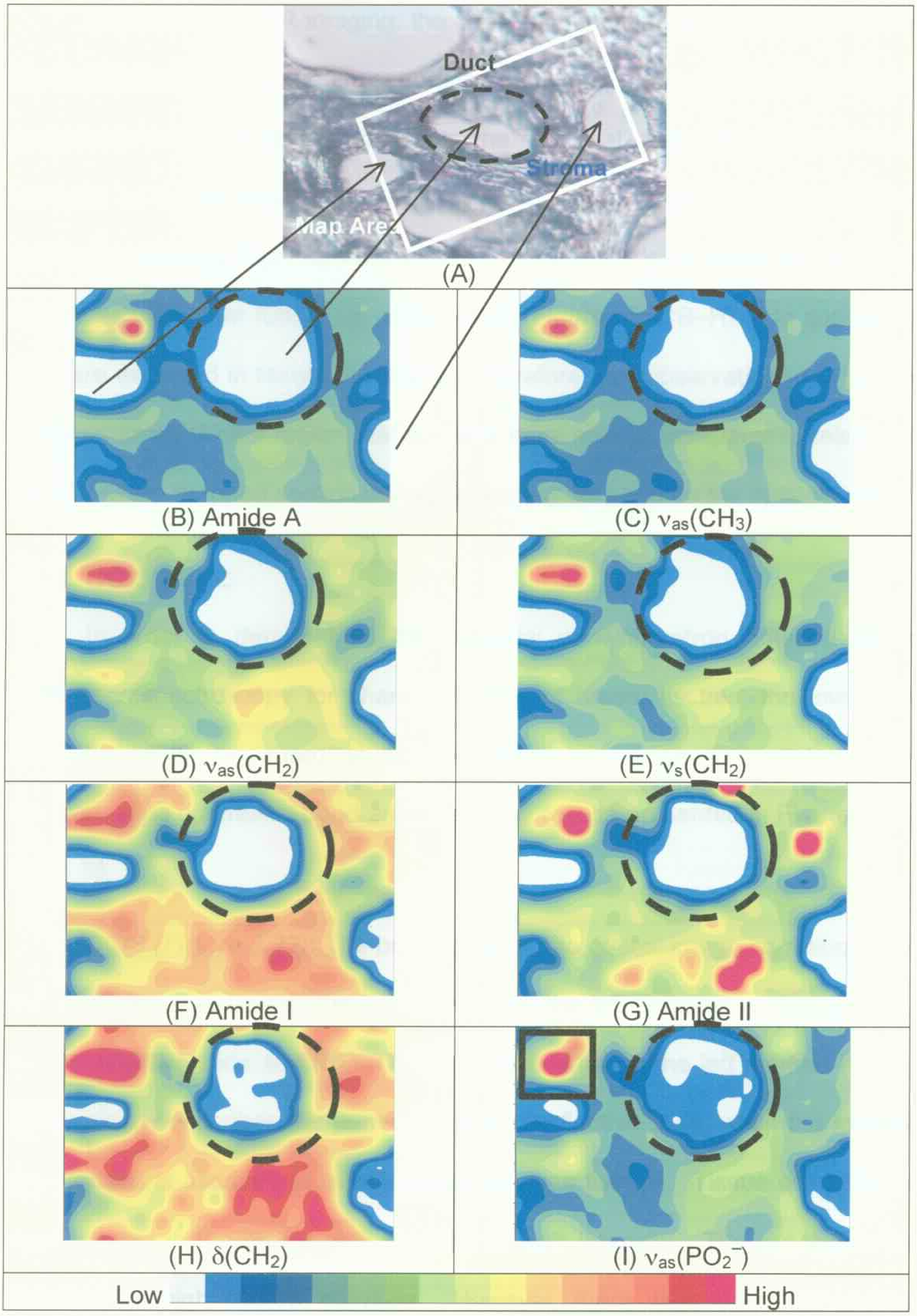
Synchrotron-radiation-based FTIR (SR-FTIR) microspectroscopy is a recent approach for tissue diagnosis [32,33,49-57] since the incident light per unit area is 100-1000 times brighter than a conventional globar source [23,37,61]. When compared to single-point microscopy, the high brightness of the collimated synchrotron source allows a smaller area (~10 × 10 μm^2) to be probed producing a spectrum with acceptable S/N [23,37,61]. Limited research has been published which uses SR-FTI microspectroscopy for characterisation of breast cancer [2,62]. Baker *et al.* utilised the advantage of the small aperture available in SR-FTIR microspectroscopy together with PCA to map successfully a micro-calcification in a malignant lesion. Previously, this was not possible due to the spatial resolution in conventional globar source FTIR microspectroscopy [2].

In the current research, three breast deparaffinised FFPE sections including those with a fibrocystic change, a low-grade cribriform DCIS and an IDC were mapped using SR-FTIR microspectroscopy at Beamline 14A, NSRRC, Taiwan. These sections were treated with xylene/histopure to minimise spectral contamination due to paraffin (Section 4.4). The aim of this study was to evaluate the potential of the technique to provide more chemical information from lesions, and to establish new areas for optimisation of protocols for future research with benchtop instruments.

5.4.1 BREAST DUCT

The SR-FTIR functional group maps from a duct of a benign fibrocystic change are presented in Figure 5.31. These maps were plotted using the band areas attributed to proteins (amides I and II), lipids (ν(CH) and δ(CH)), and nucleic acids (ν(PO_2^-)), (Figures 5.31B–I).

The large area shown in grey/dark blue and observed in all maps is assigned to the duct openings and the other smaller areas are regions that contained adipose tissues, which were dissolved during the FFPE sectioning. This observation correlated well to the optical image (Figure 5.31A). Since there are relatively lower intensities in the area marked with circle on the maps, the duct is classified as normal with reference to the adjacent H&E stained section.



(A)

(B) Amide A

(C) $\nu_{as}(CH_3)$

(D) $\nu_{as}(CH_2)$

(E) $\nu_{s}(CH_2)$

(F) Amide I

(G) Amide II

(H) $\delta(CH_2)$

(I) $\nu_{as}(PO_2^-)$

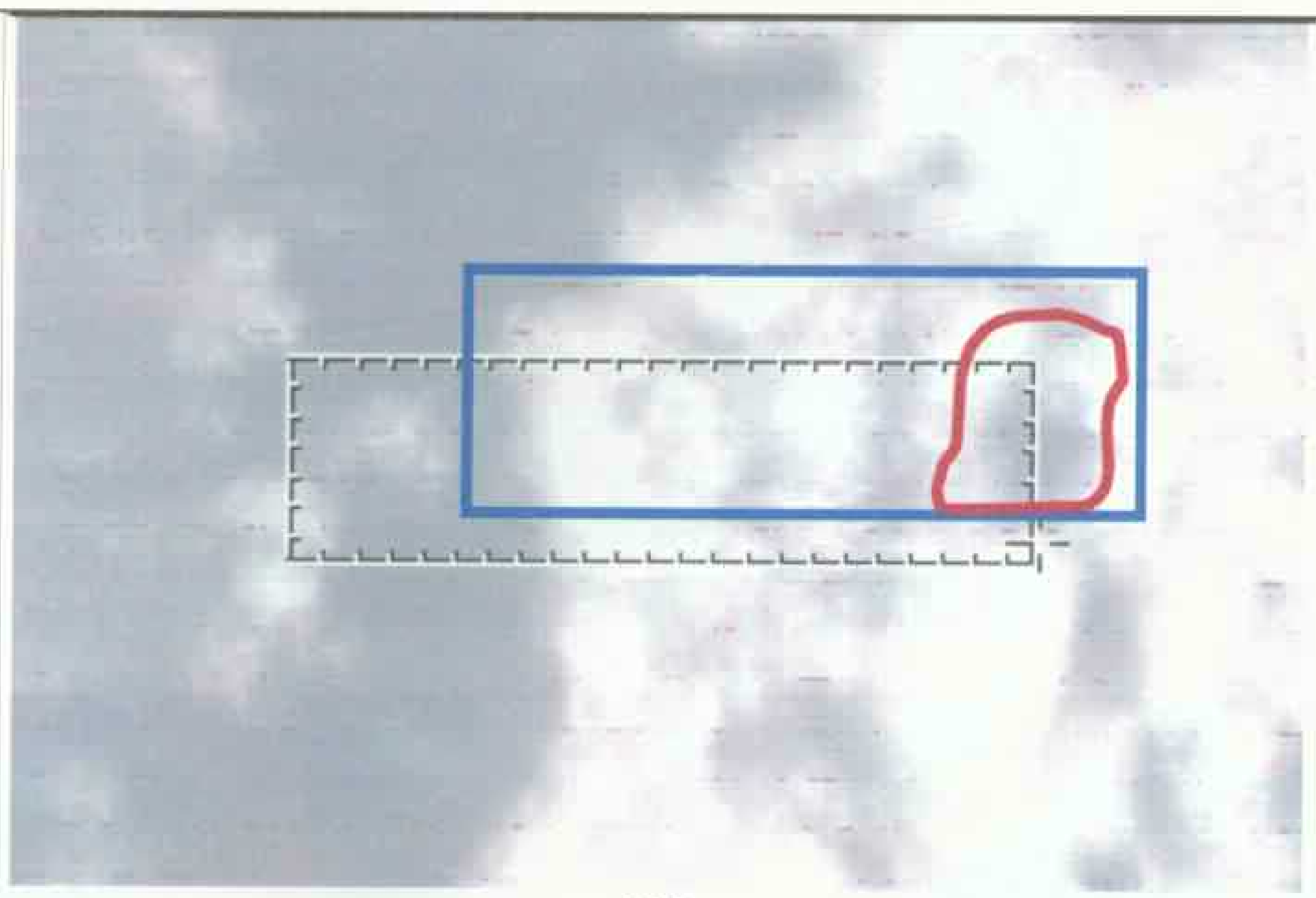
Figure 5.31. (A) Photomicrograph of a DCIS breast FFPE section and (B–I) SR-FTIR maps using the areas of diagnostic bands. Map dimension (*x*,*y*): 170 × 110 μm^2; step size: 10 × 10 μm^2; aperture: 15 × 15 μm^2; mirror velocity: 0.4747 cm s^{-1}; 32 scans; 4 cm^{-1} ; ×32 magnification and 216 total spectra. Black dotted lines: duct opening; square box: relative high intensity area. Actual collection time: 4 hr.

As with FPA FTIR imaging, the SR-FTIR functional group maps illustrate the distribution of the various tissue components. Protein is evenly distributed across the section and is attributed to the epithelial cells within the duct and collagen in the stroma. In the area marked in the nucleic acid plot (Figure 5.31I), a relatively high concentration of nucleic acids is observed and this correlates with other functional group maps (Figures 5.31B–H). No cancerous cells are expected in fibrocystic change, therefore, this observation may be due to the relatively thicker region that contains more cellular components relative to the rest of the mapped section during sectioning, as seen on the optical image.

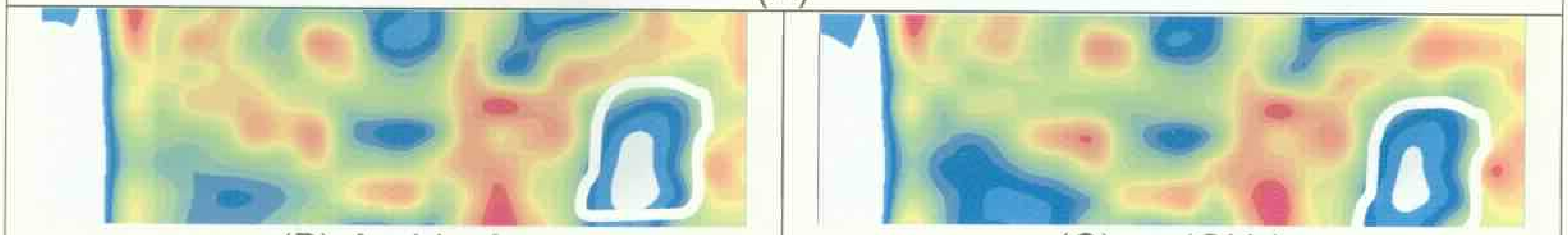
5.4.2 DCIS & IDC

In order to demonstrate the potential of synchrotron-radiation based FTIR microspectroscopy for characterisation of breast tissues, the low-grade cribriform DCIS measured at different location and an IDC section were mapped using diagnostic band areas, and the results presented in Figures 5.32 and 5.33.

***DCIS (Figure 5.32)*:** The bright area on the optical image corresponds to the DCIS tissue, whereas the darker area on the left contains no tissues, i.e., the transflective slide surface. The non-tissue area on the left marked in grey colour clearly describes the surface of the transflective slide, while the marked area (white) is ascribed to the dissolved adipose tissues. Tissue components were located in the region with medium to high intensity. The red areas are indicative of high cellular activities. However, there were only 192 spectra collected in this map (110 × 35 μm^2) using the step size of 10 × 10 μm^2 which did not provide enough statistical significant data for multivariate statistical analysis, compared with the data collected by the FPA-FTIR imaging technique.

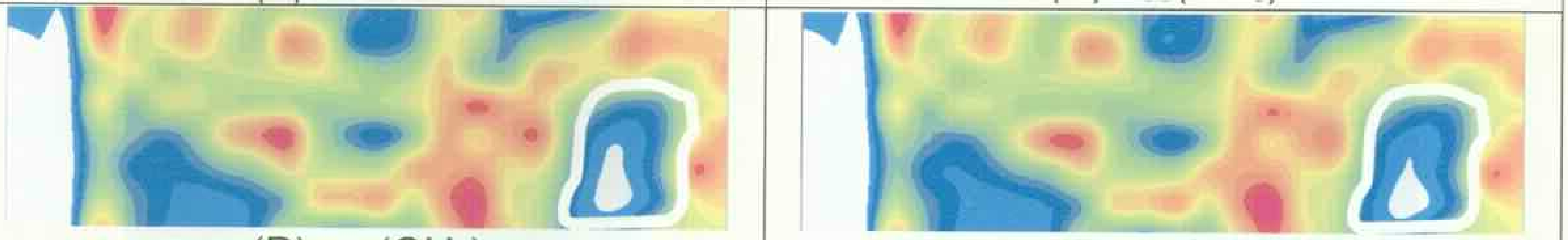


(A)



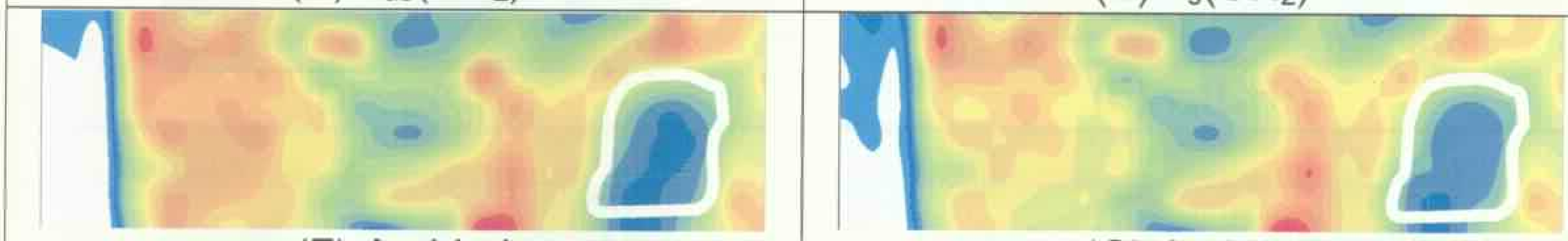
(B) Amide A

(C) $\nu_{as}(CH_3)$



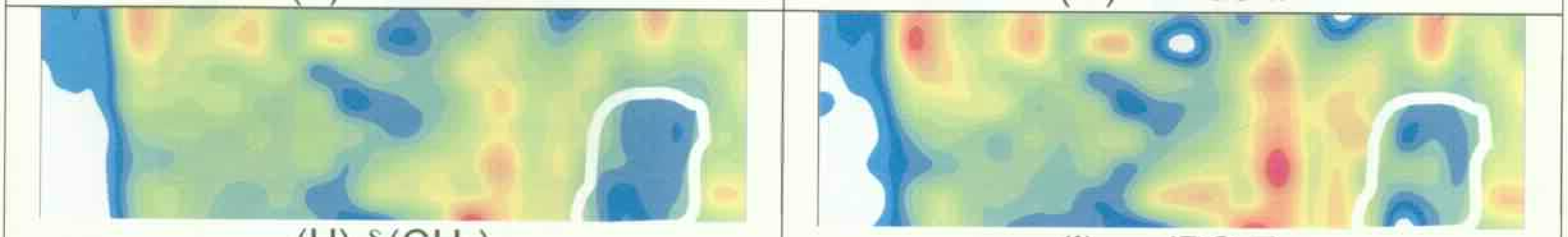
(D) $\nu_{as}(CH_2)$

(E) $\nu_s(CH_2)$



(F) Amide I

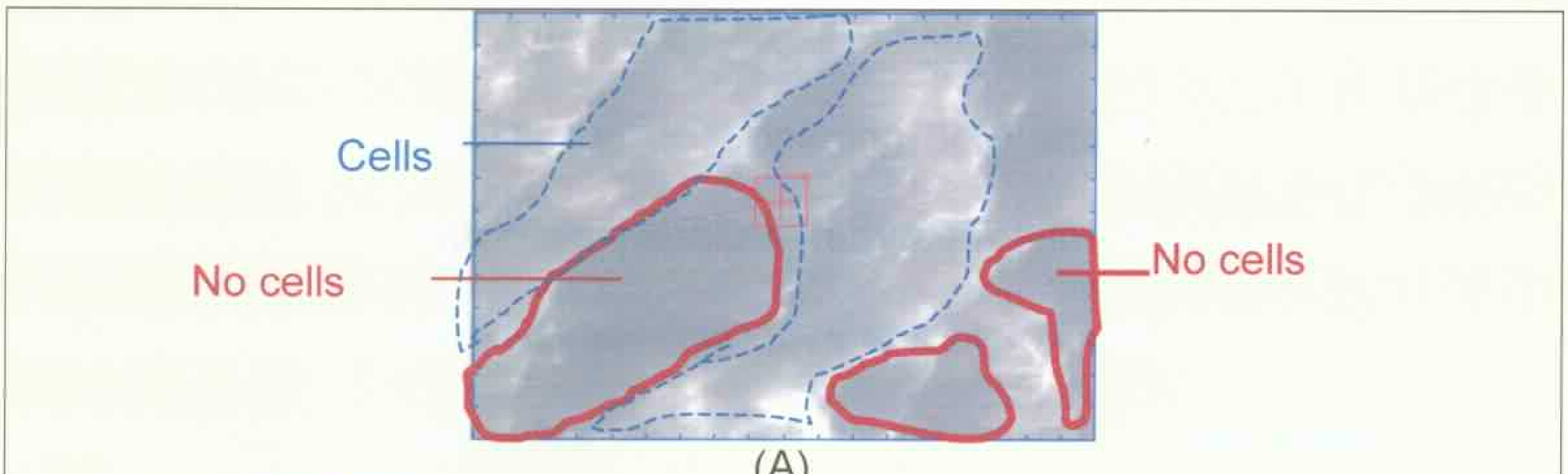
(G) Amide II



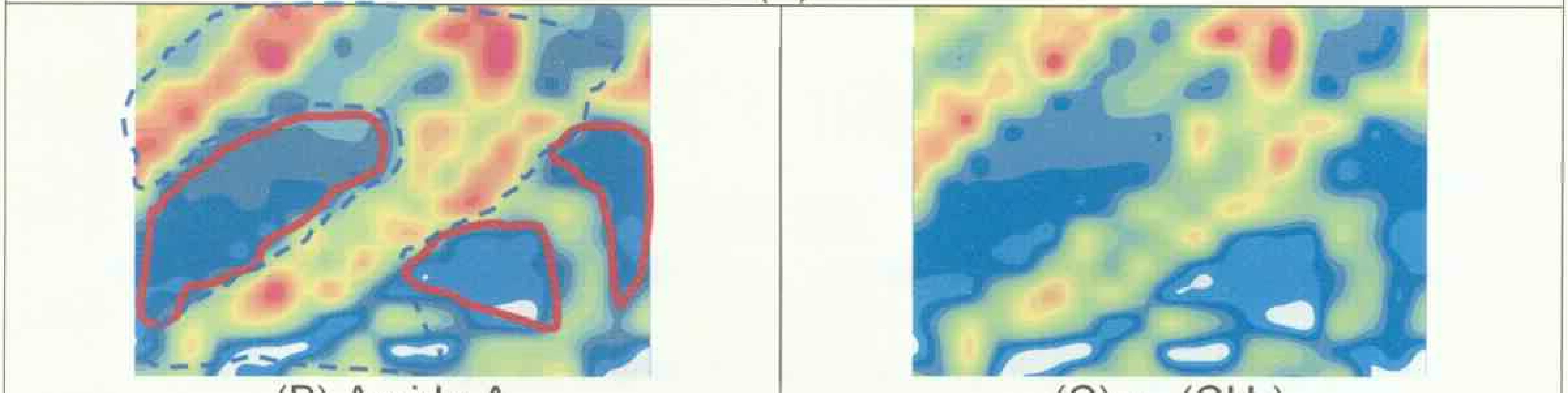
(H) $\delta(CH_2)$

(I) $\nu_{as}(PO_2^-)$

Figure 5.32. (A) Photomicrograph of an unstained DCIS breast FFPE section (deparaffinised) and (B–I) SR-FTIR functional group intensity maps. Map dimension (*x*,*y*): 110 × 35 μm^2; step size: 10 × 10 μm^2; aperture: 15 × 15 μm^2; mirror velocity: 0.4747 cm s^{-1}; 32 scans; 4 cm^{-1} res; 192 total spectra. Actual collection time: 3 hr.

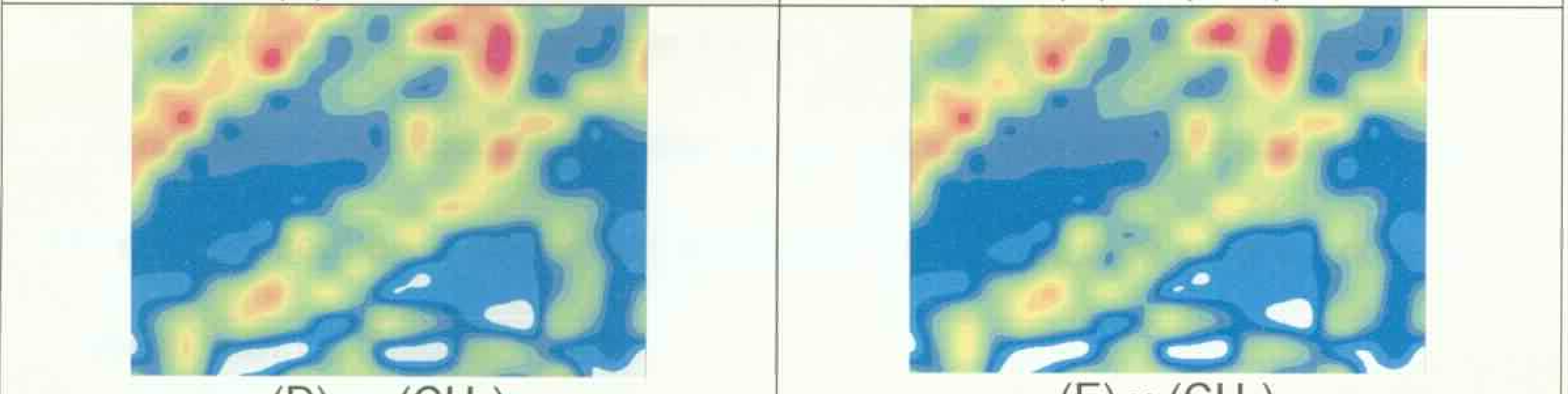


(A)



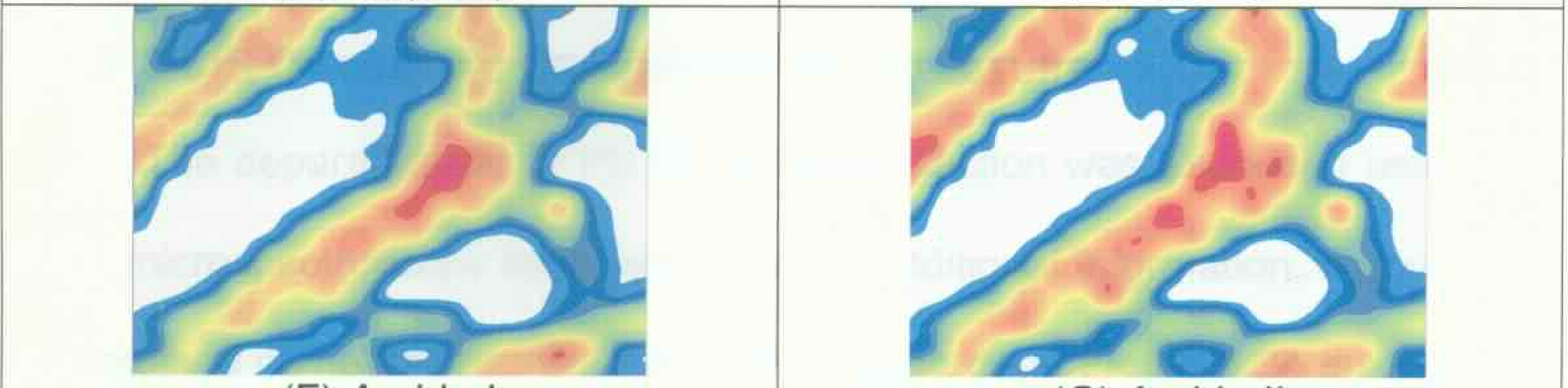
(B) Amide A

(C) $\nu_{as}(CH_3)$



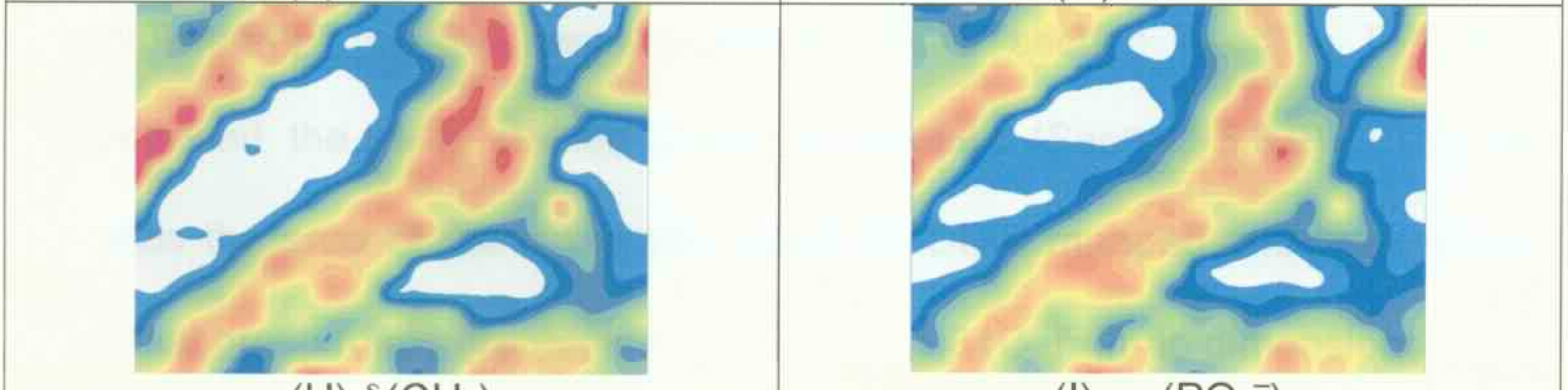
(D) $\nu_{as}(CH_2)$

(E) $\nu_s(CH_2)$



(F) Amide I

(G) Amide II



(H) $\delta(CH_2)$

(I) $\nu_{as}(PO_2^-)$

Figure 5.33. (A) Photomicrograph of an unstained malignant breast FFPE section (deparaffinised) and (B–I) SR-FTIR functional group intensity maps. Map dimension (*x*,*y*): 180 × 130 μm^2; step size:10 × 10 μm^2; aperture: 15 × 15 μm^2; mirror velocity: 0.4747 cm s^{-1}; 16 scans; 4 cm^{-1} res.; and 266 total spectra. Actual collection time: 3 hr.

IDC (Figure 5.33): The areas marked in blue or grey (low concentrations) are due to the duct opening (left) and the dissolved adipose tissues. Areas with high concentrations (red) of protein bands (amide A, I and II bands) and the nucleic acids are indicative of a higher cell density or cell nuclei at these locations than in the rest of the tissue image. These areas (marked with blue-dotted lines) were consistent with the adjacent H&E stained sections.

The tumour cells found in the lesion are indicated by the elevated concentration of diagnostic bands. As shown in Figure 5.33, the medium-to-high intensities of amide A, I and III functional group maps (Figures 5.33B, F and G), and nucleic acid maps in Figure 5.33I observed in the tissue regions are indicative of cellular proliferation when referenced to the optical image and H&E stained section. The areas displayed with dark blue or grey colours are the areas without tissues.

5.4.3 DCIS DISPLAYING CRIBRIFORM ARCHITECTURE

The deparaffinised FFPE DCIS breast section was examined using SR-FTIR microspectroscopy to determine what additional information, if any, could be obtained compared to the FPA-FTIR imaging data. Figure 5.34 presents the photomicrograph of the same unstained DCIS section used in FPA-FTIR imaging and the H&E stained DCIS cryosection (Section 5.3.2). The box indicates the area of the low-grade cribriform DCIS mapped and the resultant functional group maps are depicted in Figures 5.34C–F (map dimensions: 240 × 210 μm^2).

The functional group maps constructed from the nucleic acids bands, $\nu_{as}(PO_2^-)$ and $\nu_s(PO_2^-)$, were unable to be obtained in this particular study due to the poor S/N ratio obtained below 1300 cm^{-1}. This was due to the instrument parameters that were used at the time of this experiment, as evident in the extracted infrared spectra in Figure 5.35. Subsequent visits produced improved results as these instrument parameters were optimised.

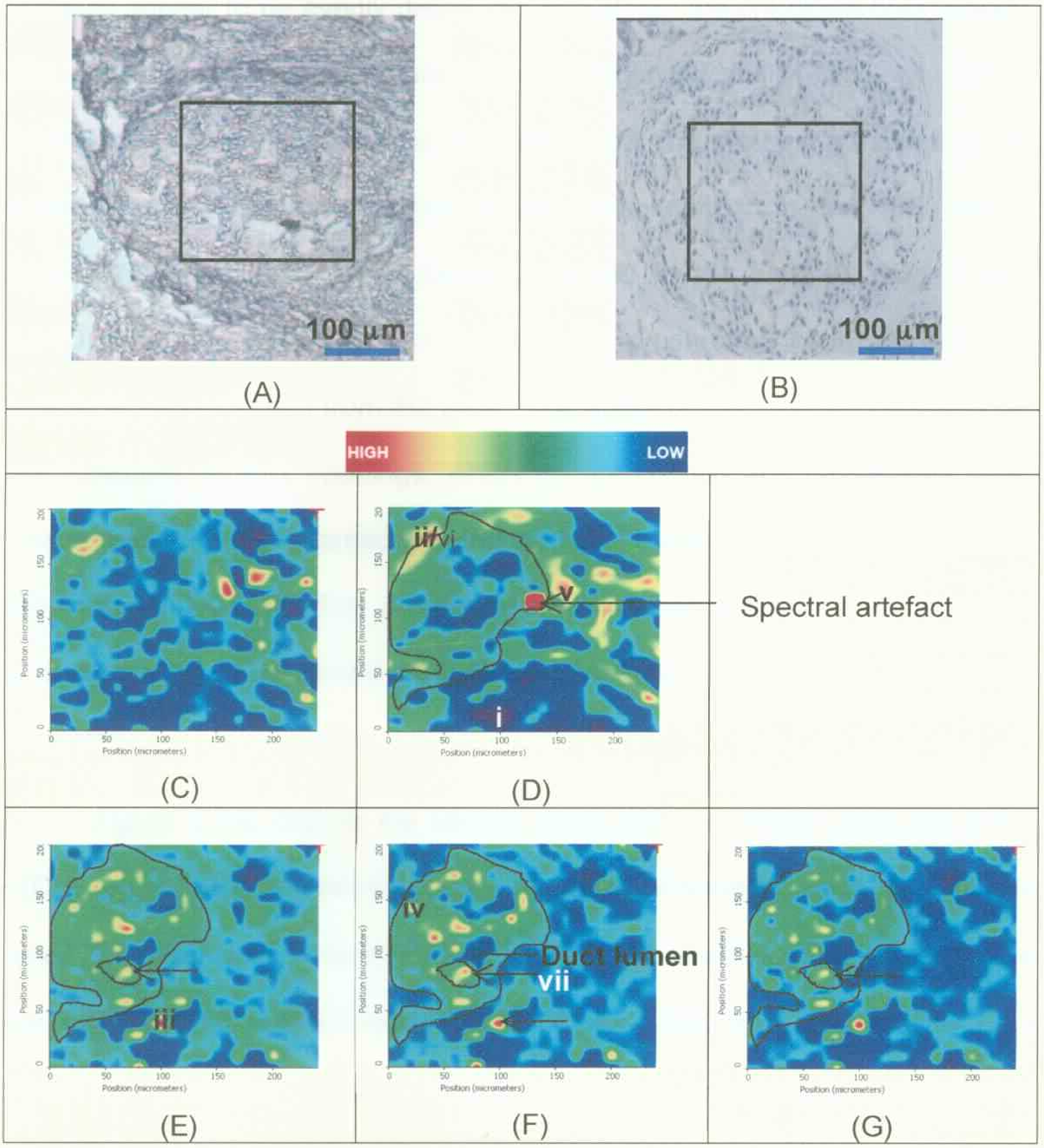


Figure 5.34. Optical images of: (A) unstained; and (B) adjacent H&E stained sections. The area imaged is marked by a black square. SR-FTIR maps of a DCIS breast duct are based on the area under the band for: (C) amide A (3630-3115 cm^{-1}); (D) amide I (1718-1588 cm^{-1}); (E) ν(CH) region (2989-2786 cm^{-1}); (F) $\nu_{as}(CH_2)$ (2943-2875 cm^{-1}); and (G) $\nu_s(CH_2)$ (2863-2818 cm^{-1}). Map dimensions (x,y): 240 × 210 μm^2 (W × H); aperture size: 15 × 15 μm^2; step size: 10 × 10 μm^2, mirror velocity: 0.4747 cm s^{-1}; 32 scans; 4 cm^{-1}; ×32 magnification; and 600 total spectra. Actual collection time: 10 hr.

Figures 5.34C and D display the functional group maps plotted using the areas under the amide A (3620–3115 cm^{-1}) and amide I (1720–1590 cm^{-1}) bands, respectively. As discussed earlier, the relative intensities of these amide bands are indicative of the cellular distribution (e.g., stroma and tumour cells). Proteins appear to be evenly distributed throughout the cribriform duct (green), although there are several areas (red) displaying high levels of cellular proliferation. This distribution is associated with the proteins attributed to the epithelial and intralobular stromal tissues, as evidenced in the H&E stained section (Figure 5.34B). The artefact, marked in the amide I map (Figure 5.34D), was due to the random noise in the synchrotron beamline and was unable to be eliminated from the data. The areas of low intensity (purple/blue) are indicative of duct openings, which do not contain any proliferative cells. These areas are interspersed with the epithelial tissue and are characteristic of the cribriform DCIS [56,63]. A representative infrared spectrum collected within a duct opening is presented in Figure 5.35A (i) and is devoid of features as expected.

Figure 5.34E depicts the infrared maps that are plotted using the entire $\nu(CH)$ region, which contains overlapping bands of proteins, lipids, and nucleic acids [31]. High lipid content results in increased intensities of the bands due to the $\nu_{as}(CH_2)$ and $\nu_s(CH_2)$ modes at 2930 and 2850 cm^{-1}, respectively [31]. Functional group maps based on the areas of these two bands are presented in Figures 5.34F and G, respectively. The spectra were collected from an area within the cribriform DCIS duct lumen, no adipose tissues would be present, therefore, the distribution of the $\nu(CH_2)$ bands observed in the marked region (Figures 5.34F and H) are attributed to the phospholipids and lipoproteins of the cell membranes of the tumour cells [8,29]. The area shown with arrows appear to be absent in the amide maps and are likely to be due to lactation products found within the duct lumen.

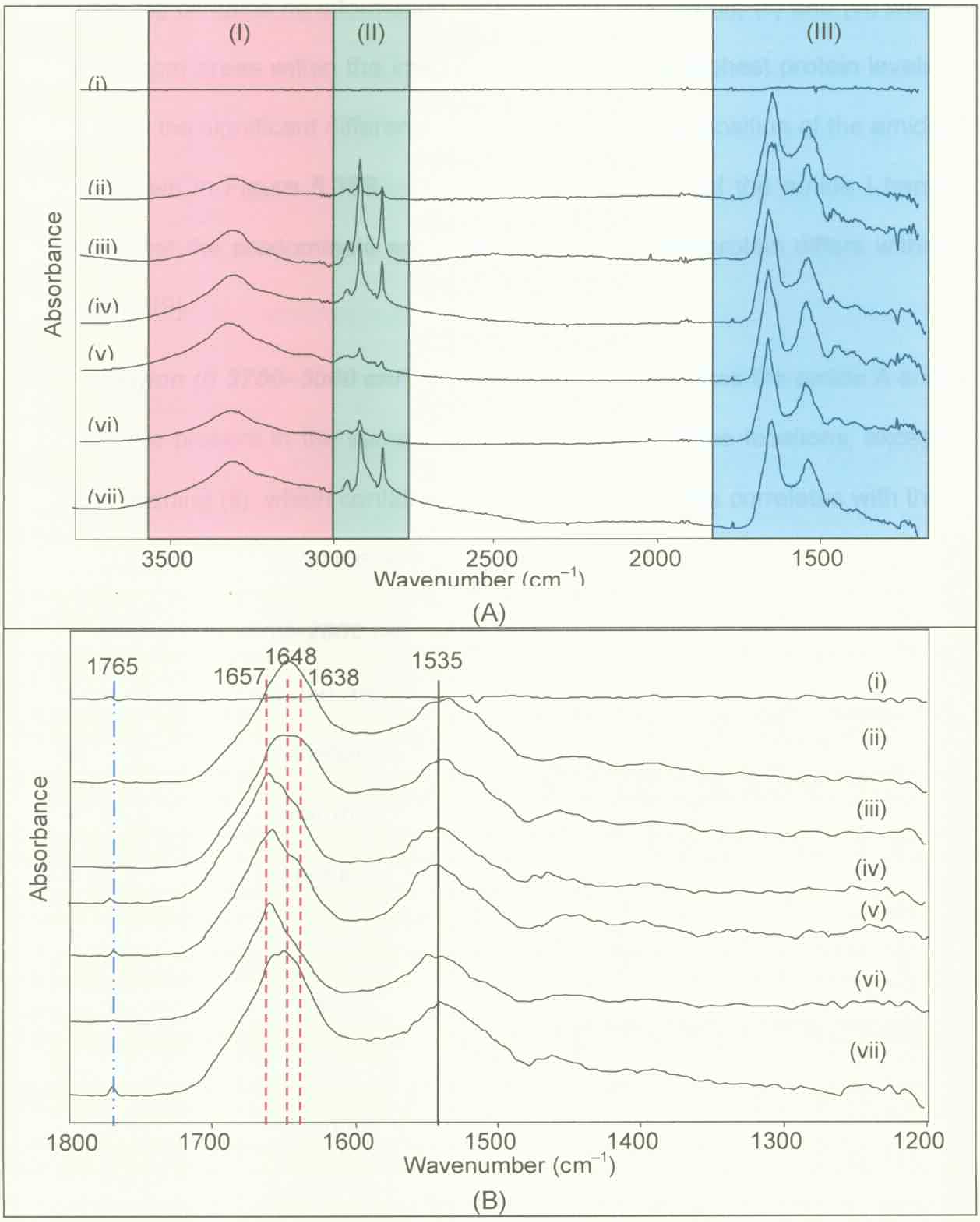


Figure 5.35. Normalised IR spectra extracted at various points (i) to (vi) in Figure 5.34 of the DCIS section using SR-FTIR microspectroscopy in the spectral region of: (A) 4000-1200 cm^{-1}; and (B) 1800-1200 cm^{-1}.

The band intensities in the FTIR spectra (i–vii) extracted from numerous locations within the DCIS section were normalised to the intensity of the amide I band and the results are presented in Figure 5.35. Spectrum (i) from the empty duct opening contains no information, as expected; spectra (ii), (v) and (vi) were extracted from areas within the image that displayed the highest protein levels. Of note are the significant differences in the lineshape and position of the amide I band shown in Figure 5.35B. The shift in the position of the amide I band suggests that the predominate secondary structure of the protein differs within the sample [9].

***Region (I) 3700–3000 cm^{-1}*:** This spectral region shows the amide A and B bands are present in the same positions in different tissue locations, except the duct opening (ii), which contains no cellular activity. This correlates with the blue/purple areas in the amide A map of Figure 5.34.

***Region (II) 3000–2800 cm^{-1}*:** The relative intensities of the $\nu(CH_2)$ bands are high in spectra (iii), (iv) and (vii). These spectra were collected from areas containing mainly of epithelial tissues and the high intensity of lipids is proposed to be due to the cell wall membranes. By contrast, the $\nu(CH_2)$ bands are of medium intensity in spectra (ii), (v) and (vi), and are low (blue) in the duct opening at position (i).

***Region (III) 1800–1300 cm^{-1}*:** The infrared spectra collected at positions (ii), (iii) and (vii) in the tissue, display amide I bands of strong intensity at 1648 cm^{-1}, while strong amide I bands centred at 1657 cm^{-1} are observed in (iv), (v) and (vi). The 1648 cm^{-1} band indicates the protein is predominately in a β-sheet conformation, whereas the band at 1657 cm^{-1} is from a protein predominately in an α-helix conformation [9]. The shifts of amide I absorption from an α-helix to β-sheet component, together with the high intensities of the bands due to $\nu(CH_2)$ found in (ii) and (iii) is likely to be due to the lipoprotein layers of epithelial cell walls within the duct lumen. In contrast, spectra (iv), (v) and (vi) may be associated with the tumour cells, as suggested by the dominant α-helix protein structure [9]. Spectrum (vii) has a medium-to-high $\nu(CH)$ band intensity but a very low amide band intensity of the contributions on the functional group maps (blue). This is indicative of cells found in the duct lumen within the cribriform duct and may be associated with milk expression, but further analyses are required to confirm this hypothesis.

5.5 CONCLUSIONS

Three infrared techniques, single-point microspectroscopy, FPA-FTIR imaging and SR-FTIR microspectroscopy, were used to characterise and differentiate breast lesions. The present study demonstrates that benign, *in-situ* carcinomas and malignant lesions in breast tumour tissue sections can be differentiated based on their IR spectra. Three discrete spectral regions were analysed, of which the fingerprint region (1700-1000 cm^{-1}) provided the diagnostic bands that were commonly used for tissue differentiation in this research and in the literature [2-28]. This is the first IR spectroscopic study undertaken which specifically looks at these type of lesions using FPA-FTIR imaging and SR-FTIR mapping (i.e., fibrocystic change, DCIS and IDC).

The visual examination of infrared spectra from breast lesions was proved to be unreliable, due to similarities of spectra from different types of lesions as shown in Section 5.2. Curve-fitting and second-derivatives were applied to the raw IR data to resolve the bands attributed to biomolecules (proteins, lipids and nucleic acids). Principal component analysis and hierarchical cluster analysis successfully separated spectra obtained from malignant lesions and those obtained from benign lesions using single-point microspectroscopy over the spectral region of 1700-1000 cm^{-1}. PCA clearly distinguished the spectra of fibrocystic changes from the spectra of benign fibroadenomas and intraduct papillomas. Although the spectra of IDC and ILC sections were not clearly separated using the raw data, PCA using second-derivatives of the spectra distinguished both IDC and ILC lesions from each other. Two pre-classified IDC unstained lesions were clustered as DCIS, which could be associated with the co-existence of the morphological features from DCIS (IR measurement taken) and IDC. Similarly, HCA differentiated benign and malignant breast lesions using the combination of spectral regions. Two breast lesions without prior knowledge to the histopathological outcomes were classified as fibrocystic changes using HCA, which were then correlated with the histopathological results. The IR spectra of two pre-classified malignant IDC tissue sections were clustered using HCA within the DCIS (D) sub-group. These two lesions were subsequently stained after IR analysis and histopathology results confirmed that these lesions were, in fact, high-grade DCIS in which cell proliferation was in the process of invading through the basement membrane, as evident from the adjacent H&E stained section. The HCA classification and subsequent histopathology examination overruled the original classification of being malignant IDC.

Spectral classification of benign, *in-situ* carcinomas and malignant lesions using PCA and HCA were achieved with overall sensitivities of 100/100 (PCA/HCA), 100/78 and 92/91%, and specificities of 100/94, 96/96 and 100/100%, respectively. The results were comparable with previous research using ^{1}H MRS which distinguished malignant tissues from benign lesions with sensitivity and specificity of 95 and 96 per cent, respectively [51].

In this research, various spectroscopic techniques were used to obtain spectral information that could characterise and differentiate breast lesions. It was deduced that by combining the focal-plane array FTIR imaging with multivariate statistical analysis, in particular k-means clustering, information that could be easily interpreted by spectroscopists and pathologists was obtained. This methodology was based on the collection of multiple spectra simultaneously (4096 spectra) using a focal-plane array in a relatively short time (<5 min), compared to the single-point mapping in both conventional globar-source and synchrotron-source FTIR mapping (> 2 hr). However, there is limited published research on IR imaging of breast lesions [10,11,42] and there is a need to improve current breast cancer diagnostic research using FPA-FTIR imaging. Prior to any multivariate statistical analysis (MSA) being applied to the imaging data, the tissue composition was examined using the adjacent H&E stained section in order to determine the optimal clusters used for the specific tissue section.

Three deparaffinised FFPE breast sections, including a normal breast duct, a low-grade cribriform DCIS and an IDC, were imaged and analysed using KMC, FCM and/or PCA. Five clusters were proved to best identify the major tissue components within the lesions. These components included epithelial cells, tumour cells, stromal tissues, duct openings/dissolved adipose tissues and basement membranes, which exhibited significant spectral features particularly in the fingerprint region. As shown by the spectra collected using single-point FTIR microspectroscopy, each breast disease condition had characteristic morphological and spectral features that could be distinguished using MSA. The MSA techniques examined the spectral similarities from each recorded pixel element of the IR imaging hyperspectral cube and the resulting cluster maps were assigned to specific tissue components (colours) that were well-correlated to the histopathology. Insufficient cluster numbers (<3) used in MSA distinguished the major tissue components, i.e., cells and non-cellular components, but the extracted mean spectra clearly demonstrated that bands due to biomolecules were still present in the cluster of non-cellular components (Figure 5.21), while an excessive number (>5) of clusters could produce ambiguous images when there are more clusters than know histopathological features.

Synchrotron-radiation source FTIR microspectroscopy was used to investigate the diagnostic potential for cancer research. The advantage of SR-FTIR microspectroscopy is that the incident light per unit area is 100-1000 times brighter than a conventional globar source, which produces a spectrum with acceptable S/N ratio in a small aperture (~10 × 10 μm^2) [23,37,61]. Currently, there are two published studies on breast cancer using SR-FTIR microspectroscopy [2,62]. Baker *et al.* utilised the advantage of the small aperture available in SR-FTIR microspectroscopy together with PCA to map successfully a micro-calcification in a malignant lesion [2]. Liu, *et al.* examined the spectral different breast diseases using single-point SR-FTIR microspectroscopy [62]. In the current study, three deparaffinised breast FFPE sections, a fibrocystic change, a low-grade cribriform DCIS and an IDC, were mapped. Although, relatively weak signals were observed below 1300 cm^{-1} in all tissue spectra, the functional group maps were highly correlated to the tissue components in H&E stained sections.

With an overall sensitivity of 97.3% and specificity of 98.5% for breast lesion classification (PCA) using the mid-infrared region, infrared microspectroscopy is a proven supportive technique to the gold standard of histopathology and magnetic resonance spectroscopy for cancer diagnosis. It is envisaged that further development of instrumentation and computer software will eventually mean infrared microspectroscopy will be a stand-alone technique for cancer diagnosis.

5.6 REFERENCES

[1] D. Naumann, *Appl. Spectrosc. Rev.* 36 (2001) 239-298.

[2] R. N. Baker, J. Smith, K. D. Rogers, N. Stone, *Proc. SPIE-Int. Soc. Opt. Eng.* 6093 (2006) 60930X/60931-60930X/60911.

[3] K. Bambery, B. Wood, M. A. Quinn, D. McNaughton, *Aust. J. Chem.* 57 (2004) 1139-1143.

[4] L. Chiriboga, H. Yee, M. Diem, *Appl. Spectrosc.* 54 (2000) 480-485.

[5] L. Chiriboga, H. Yee, M. Diem, *Appl. Spectrosc.* 54 (2000) 1-8.

[6] M. Diem, J. Bargonetti, T. Gopen, S. Boydston-White *Method and Apparatus for Characterization of Biological Entities Such as Cells Using Ir Spectra*. Patent: 0043775, 2000.

[7] M. Diem, L. Chiriboga, H. Yee, *Biopolymers* 57 (2000) 282-290.

[8] R. Dukor, *Vibrational Spectroscopy in the Detection of Cancer in Handbook of Vibrational Spectroscopy.* 1st ed., John Wiley & Sons, Chichester, 2002, Vol. 5, 3335-3361.

[9] R. Eckel, H. Huo, H.-W. Guan, X. Hu, X. Che, W.-D. Huang, *Vib. Spectrosc.* 27 (2001) 165-173.

[10] H. Fabian, P. Lasch, M. Boese, W. Haensch, *Biospectroscopy* 67 (2002) 354-357.

[11] H. Fabian, P. Lasch, M. Boese, W. Haensch, *J. Mol. Struct.* 661-662 (2003) 411-417.

[12] H. Fabian, W. Mäntele, *Infrared Spectroscopy of Proteins in Handbook of Vibrational Spectroscopy.* 1st ed., John Wiley & Sons, Chichester, 2002, Vol. 5, 3399-3425.

[13] E. O. Faolain, M. B. Hunter, J. M. Byrne, P. Kelehan, H. J. Byrne, F. M. Lyng, *Proc. SPIE-Int. Soc. Opt. Eng.* 5826 (2005) 25-36.

[14] T. Gao, Y. Ci, *Anal. Commun.* 36 (1999) 341-343.

[15] T. Gao, J. Feng, Y. Ci, *J. Eur. Soc. Anal. Cell. Path.* 18 (1999) 87-93.

[16] M. Jackson, *Microsc. Microanal.* 10 (Suppl. 2) (2004) 1292-1293.

[17] M. Jackson, L. Choo, P. H. Watson, W. C. Halliday, H. H. Mantsch, *Biochim. Biophys. Acta* 1270 (1995) 1-6.

[18] M. Jackson, K. Kim, J. Tetteh, J. R. Mansfield, B. Dolenko, R. L. Somorjai, F. W. Orr, P. H. Watson, H. H. Mantsch, *Proc. SPIE-Int. Soc. Opt. Eng.* 3257 (1998) 24-34.

[19] M. Jackson, J. R. Mansfield, B. Dolenko, R. L. Somorjai, H. H. Mantsch, *Cancer Detect. Prev.* 23 (1999) 245-253.

[20] J. Kneipp, *Vibratioanl Spectra from Tissues, Cells and Sub-Cellular Strutures.* In 5th Australian Conference on Vibrational Spectroscopy, ACOVS6 Organising Committee: University of Sydney, 2005.

[21] P. Lasch, W. Haensch, D. Naumann, M. Diem, *Biochem. Biophys. Acta* 1688 (2004) 176-186.

[22] P. Lasch, D. Naumann, *Cell. Mol. Biol.* 44 (1998) 189-202.

[23] L. Miller, P. Dumas, *Biochem. Biophys. Acta* 1758 (2006) 846-857.

[24] W. Petrich, *Appl. Spectrosc. Rev.* 36 (2001) 181-237.

[25] S. Shen, B. Liu, X. Ma, Z. Song, Q. Li, *Guangpuxue Yu Guangpu Fenxi* 20 (2000) 28-30.

[26] B. Wood, L. Chiriboga, H. Yee, M. A. Quinn, D. McNaughton, M. Diem, *Gynecol. Oncol.* 93 (2004) 59-68.

[27] B. R. Wood, M. A. Quinn, B. Tait, M. Ashdown, T. Hislop, M. Romeo, D. McNaughton, *Biospectroscopy* 4 (1998) 75-91.

[28] L. Zhang, G. Small, A. Haka, L. Kidder, E. Lewis, *Appl. Spectrosc.* 57 (2003) 14-22.

[29] M. Jackson, H. H. Mantsch, *Pathology by Infrared and Raman Spectroscopy in Handbook of Vibrational Spectroscopy.* John Wiley & Sons, Chichester, 2002, Vol. 5, 3227-3245.

[30] P. Lasch, D. Naumann, *Biochim. Biophys. Acta* 1758 (2006) 814-829.

[31] G. Socrates, *Infrared and Raman Characteristic Group Frequencies: Tables and Charts.* 3rd ed., John Wiley & Sons Ltd., Chichester, 2001.

[32] The Susan G. Komen Breast Cancer Foundation (USA), http://www.komen.org 2005.

[33] G. McKee, *Cytopathology of the Breast.* 1st ed., Oxford University Press, Boston, 2002.

[34] H. Fabian, R. Wessel, M. Jackson, A. Schwartz, P. Lasch, I. Fichtner, H. H. Mantsch, D. Naumann, *Proc. SPIE-Int. Soc. Opt. Eng.* 3257 (1998) 13-23.

[35] M. Meurens, J. Wallon, J. Tong, H. Noeel, J. Haot, *Vib. Spectrosc.* 10 (1996) 341-346.

[36] H. Fabian, M. Jackson, L. Murphy, P. H. Watson, I. Fichtner, H. H. Mantsch, *Biospectroscopy* 1 (1995) 37-45.

[37] L. Miller, *J. Biol. Phys.* 29 (2003) 219-230.

[38] M. Diem, M. Romeo, *Vib. Spectrosc.* 38 (2005) 115-119.

[39] M. Diem, M. Romeo, C. Matthaus, M. Miljkovic, L. Miller, P. Lasch, *Infrared Phys. Techn.* 45 (2004) 331-338.

[40] M. Jackson, J. R. Mansfield, B. Dolenko, R. L. Somorjai, H. H. Mantsch, P. H. Watson, *Cancer Detect. Prev.* 23 (1999) 245-253.

[41] C. P. Schultz, K.-Z. Liu, E. A. Salamon, K. T. Riese, H. H. Mantsch, *J. Mol. Struct.* 480-481 (1999) 369-377.

[42] H. Fabian, N. Thi, M. Eiden, P. Lasch, J. Schmitt, D. Naumann, *Biochem. Biophys. Acta* 1758 (2006) 874-882.

[43] N. Pleshko, A. LBoskey, R. Mendelsohn, *Calcif. Tissue Int.* 51 (1992) 72-77.

[44] E. Faolain, M. Hunter, J. Byrne, P. Kelehan, M. McNamara, H. Byrne, F. Lyng, *Vib. Spectrosc.* 38 (2005) 121-127.

[45] T. Gao, J. Li, Y. Ci, *Fenxi Kexue Xuebao* 16 (2000) 353-357.

[46] M. Jackson, B. Ramjiawan, M. Hewko, H. H. Mantsch, *Cell. Mol. Biol.* 44 (1998) 89-98.

[47] L. M. McIntosh, J. R. Mansfield, A. N. Crowson, H. H. Mantsch, M. Jackson, *Biospectroscopy* 5 (1999) 265-275.

[48] R. Blamey, A. Evans, I. Ellis, R. Wilson, *Atlas of Breast Cancer.* 1st ed., Merit Publishing International, Basingstoke, 1994.

[49] E. Frykberg, *Breast J.* 5 (1999) 287-350.

[50] Wikipedia, Sensitivity (Tests). http://en.wikipedia.org/wiki/Sensitivity_%28tests%29 USA, 2007.

[51] C. Lean, R. Somorjai, I. Smith, P. Russell, C. Mountford, *Ann. Rev. NMR. S.* 48 (2002) 71-111.

[52] J. Harris, *Diseases of the Breast.* 3rd ed., Lippincott-Raven, Philadelphia, 2004.

[53] M. Romeo, M. Diem, *Vib. Spectrosc.* 38 (2005) 129-132.

[54] Cytospec, *Software: Cytospec, Version 1.3*; 2006.

[55] G. Leonard, S. Swain, *J. Nat. Cancer. I.* 96 (2004) 906-920.

[56] H. Burstein, K. Polyak, J. Wong, S. Lester, C. Kaelin, *N. Engl. J. Med.* 350 (2004) 1430-1441.

[57] K. A. Skinner, *The Clinical Management of Ductal Carcinoma in-Situ, Lobular Carcinoma in-Situ and Atypical Hyperplasia of the Breast.* 1st ed., National Breast Cancer Centre, Sydney, 2003.

[58] Y. Ci, T. Gao, J. Feng, Z. Guo, *Appl. Spectrosc.* 53 (1999) 312-315.

[59] S. Samiani, M. Ludvikova, G. Tomasic, S. Bianchi, G. A., V. Eusebi, *Virchows. Arch.* 434 (1999) 227-234.

[60] R. Brereton, *Chemometrics: Data Analysis for the Laboratory and Chemical Plant.* 1st ed., John Wiley & Sons, Chichester, 2003.

[61] L. Miller, R. Smith, *Vib. Spectrosc.* 38 (2005) 237-240.

[62] C. Liu, Y. Zhang, X. Yan, X. Zhang, C. Li, W. Yang, D. Shi, *J. Lumin.* 119-120 (2006) 132-136.

[63] Imaginis, Ductal Carcinoma *in-Situ*.
http://www.imaginis.com/breasthealth/dcis.asp 2005.

6. CHARACTERISATION OF BREAST LESIONS USING RAMAN SPECTROSCOPY

6.1 SAMPLE OPTIMISATION

The sample presentation methods used in Raman spectroscopy are similar to those used in histopathology and IR spectroscopy. Sample optimisation for Raman spectroscopy includes: the selection of laser excitation(s); the evaluation of the diagnostic utility of cryosections and FFPE sections for pathology; the effect of tissue dehydration; and the selection of optimal tissue thickness and mounting substrate.

6.1.1 SELECTION OF LASER EXCITATION

Figure 6.1 compares the Raman spectra of 20-μm-thick breast cryosections mounted on CaF_2 windows acquired using various laser excitations. The aim of this experiment was to select the optimal laser excitation(s) for further development of this research.

All spectra contain bands attributable to biomolecules (proteins, lipids, β-carotene, etc.). A full assignment of the Raman bands observed from the tissues is given in Section 6.2. Visible laser excitations of 647 and 752 nm (Figures 6.1B–C) exhibit a strong fluorescence background, whereas both near-infrared (NIR) excitations at 830- and 1064-nm do not exhibit such fluorescence. These NIR excitations have been reported previously as the preferred laser excitations for diagnostics due to the information-rich Raman spectra they produce [1-14]. In this study, three laser excitations: 514, 830 and 1064 nm, were used for sample optimisation, differentiation and characterisation of breast lesions.

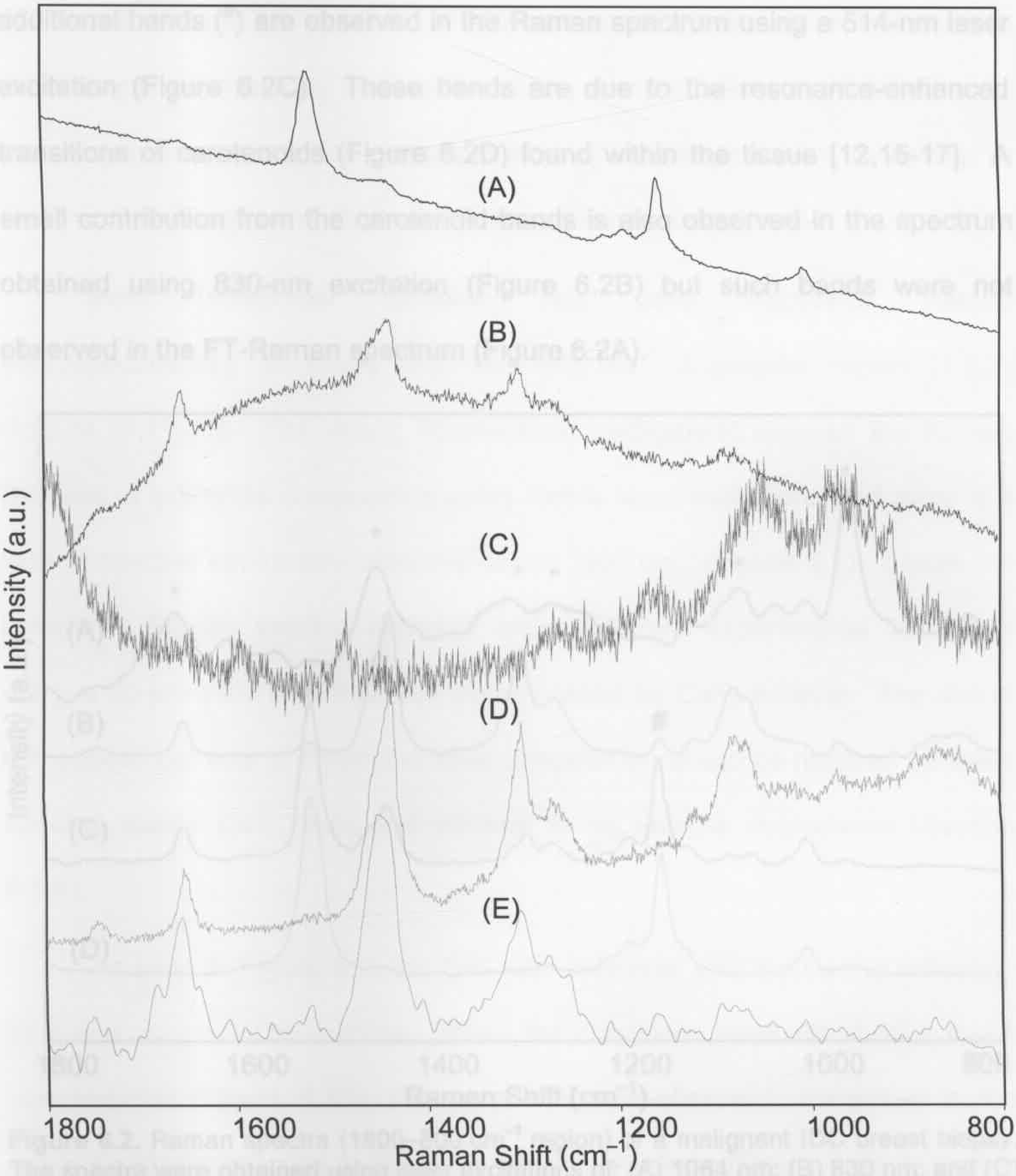


Figure 6.1. Selection of laser excitation(s) for breast tissue analysis. Raman spectra in the 1800–800 cm^{-1} spectral region of 20-μm-thick breast cryosections (regardless of tissue conditions) mounted on CaF_2 windows using laser excitations of: (A) 514 nm; (B) 647 nm; (C) 752 nm; (D) 830 nm; and fresh tissue using a laser excitation of (E) 1064 nm. Spectra are not corrected for instrument response and raw data are displayed to illustrate varying baselines and effect of fluorescence. Spectra are offset for clarity.

Figure 6.2 shows spectra from an IDC breast cryosection obtained using 514-, 830- and 1064-nm excitations. Both 830- and 1064-nm laser excitations (Figures 6.2A and B) produced spectra containing information from biological components including amide I and III bands and lipid $\delta(CH_2)$ band (*), whereas additional bands (#) are observed in the Raman spectrum using a 514-nm laser excitation (Figure 6.2C). These bands are due to the resonance-enhanced transitions of carotenoids (Figure 6.2D) found within the tissue [12,15-17]. A small contribution from the carotenoid bands is also observed in the spectrum obtained using 830-nm excitation (Figure 6.2B) but such bands were not observed in the FT-Raman spectrum (Figure 6.2A).

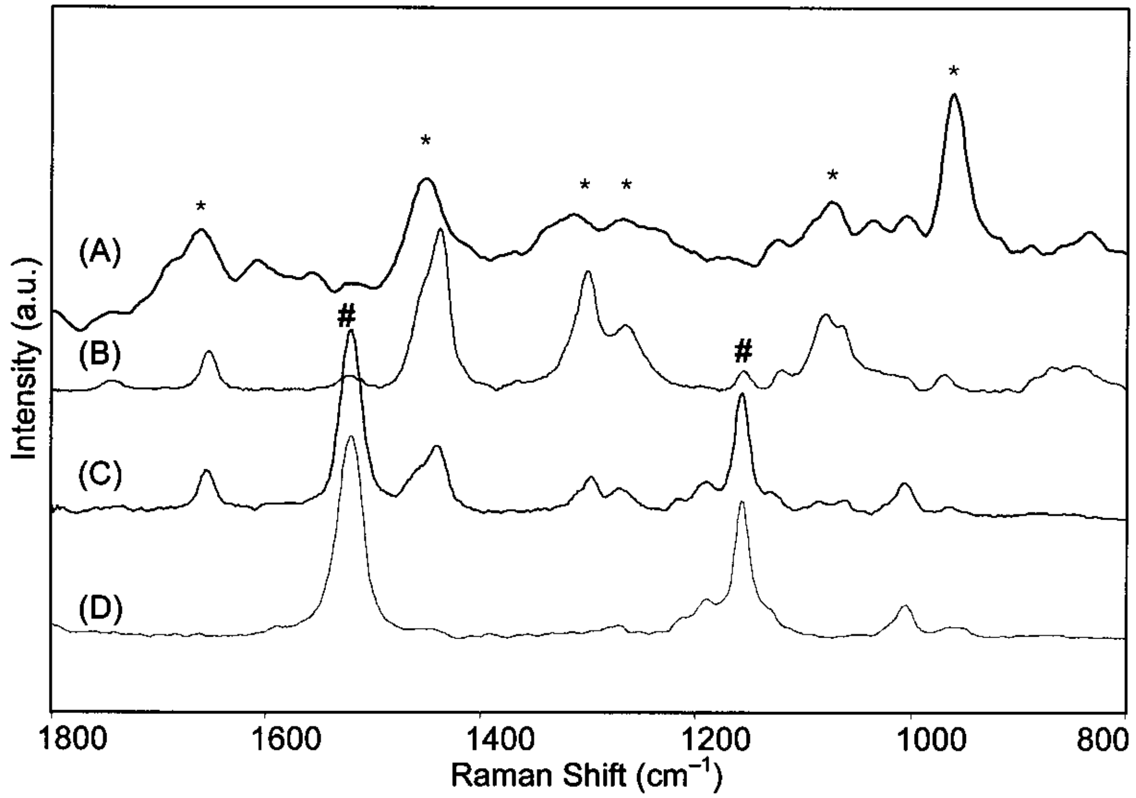


Figure 6.2. Raman spectra (1800–800 cm^{-1} region) of a malignant IDC breast biopsy. The spectra were obtained using laser excitations of: (A) 1064 nm; (B) 830 nm; and (C) 514 nm; and (D) a standard β-carotene sample obtained using laser excitation of 514-nm. Experimental conditions: (A) 'bulk' biopsy macroscopic analysis: 200 mW at sample; 256 scans; 4 cm^{-1}; (B and C) cryosection microscopic analyses: 2 mW at sample; ×20 magnification; 30 s exposure time and 5 accumulations; and (D) 0.2 mW at sample; 10 s; and 1 accumulation. Bands assigned to the biological components are marked with (*) and bands are attributed to carotenoids are marked with (#). Spectra are baseline corrected and are offset for clarity.

6.1.2 OPTIMISATION OF EXPERIMENTAL CONDITIONS FOR MICROSPECTROSCOPY

Once the laser excitations were selected, the next step was to optimise the other experimental conditions, the details of which are summarised in

Chapter 2, Tables 2.5 and 2.6. FT-Raman spectra were measured using a macroscopic approach by placing the biopsy sample (either in PBS/D_2O or in formalin) into an aluminium sample holder, while samples were measured in 'bulk' or in sections using dispersive Raman microspectroscopy.

Most biomedical research has used dispersive Raman microspectroscopy to probe the 1800–800 cm^{-1} fingerprint region [1-3,7-9,12,14,15,17-31]. The strong fluorescence background masked the Raman features of biological components using visible laser excitations and there is a lack of spectral information obtained above 1800 cm^{-1} (Figure 6.1). Figure 6.3 compares Raman spectra obtained under different experimental conditions using a 20-μm-thick breast cryosection mounted on CaF_2 window. The aim of this experiment was to determine what compromise would be required between spectral quality (S/N ratio) and minimal to no sample degradation (Section 6.1.4).

As seen in Figure 6.3, the S/N ratio improves with increasing collection time and accumulation number. When the conditions were set at 10 s and 1 accumulation (Figure 6.3A), the worst S/N was observed, compared to the same location measured at 30 s with 10 accumulations (Figure 6.3F, collection time: ~400 s). However, the longer exposure times are not practical as the tissue is damaged by the high laser powers (2 mW at the sample in a 1-μm laser spot size for the dispersive Raman microscope). For subsequent experiments, a 30-s collection time with 5 accumulations (Figure 6.3E) was selected for further microspectroscopic measurements as there are no signs of spectral changes (band shifting and intensities) occurred in the spectra.

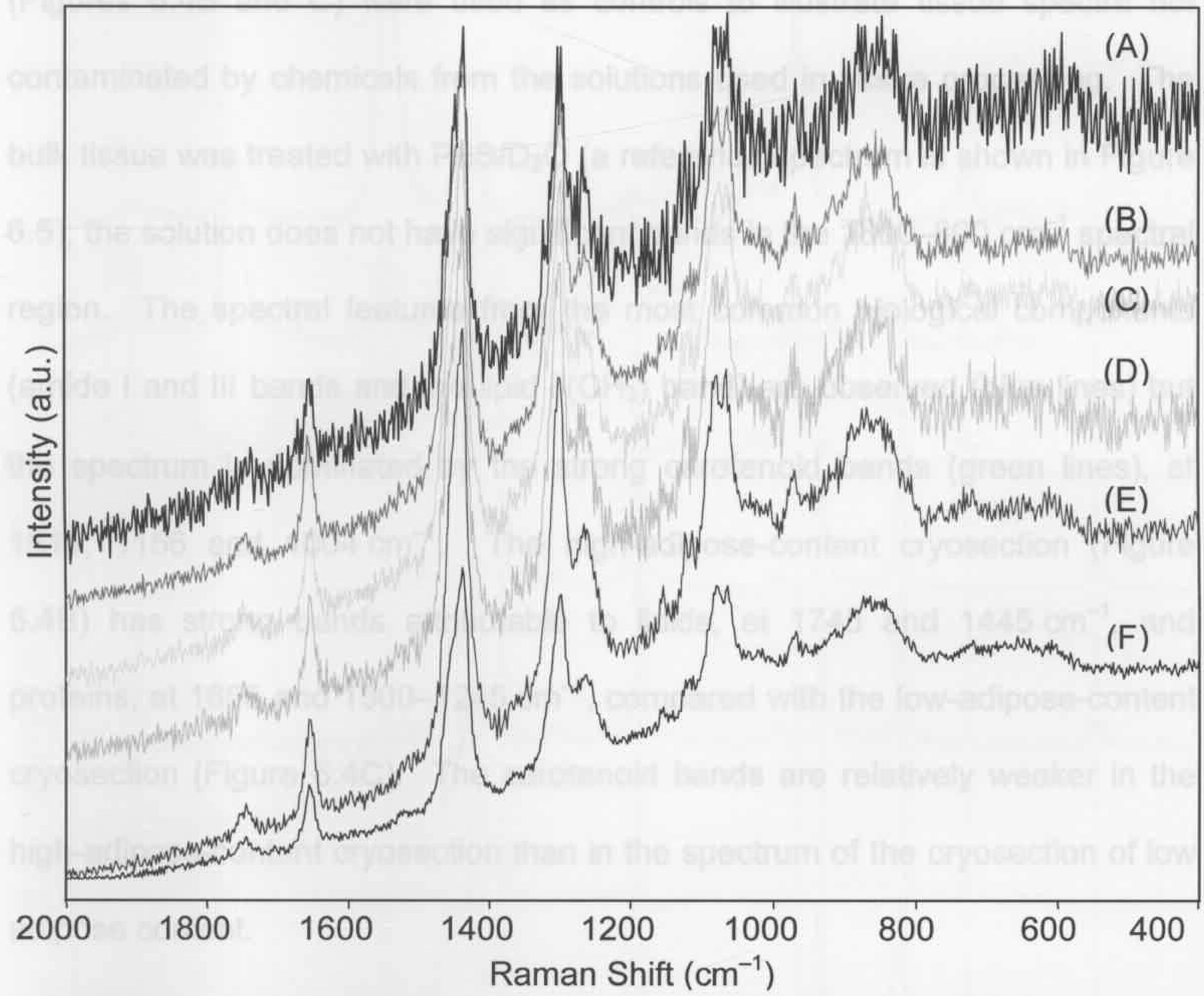


Figure 6.3. Comparison of Raman spectra of a 20-µm-thick malignant IDC breast cryosection collected with various experimental settings with respect to collection time and number of accumulations. (A) 10 s and 1 acc.; (B) 10 s and 5 acc.; (C) 10 s and 10 acc.; (D) 30 s and 1 acc.; (E) 30 s and 5 acc.; and (F) 30 s and 10 acc. Other experimental conditions: λ = 830 nm, 2 mW at sample and ×20 magnification. Spectra are offset for clarity.

6.1.3 TISSUE PROCESSING

Raman spectra were collected using 514-nm laser excitation (Figure 6.4) from a breast specimen at each stage of tissue preparation and were processed as described in Chapter 2, Section 2.2 (Figure 2.1). The aim of this experiment was to identify the presence of exogenous chemicals remaining from tissue processing, especially during FFPE sectioning. Such contamination was shown to be problematic in IR spectra of FFPE sections, as discussed in Chapter 4, Section 4.4.

The spectra of a fresh tissue biopsy (Figure 6.4A) and two cryosections

(Figures 6.4B and C) were used as controls to illustrate tissue spectra not contaminated by chemicals from the solutions used in tissue processing. The bulk tissue was treated with PBS/D_2O (a reference spectrum is shown in Figure 6.5); the solution does not have significant bands in the 1800–800 cm^{-1} spectral region. The spectral features from the most common biological components (amide I and III bands and the lipid $\delta(CH_2)$ band) are observed (blue lines) but the spectrum is dominated by the strong carotenoid bands (green lines), at 1514, 1156 and 1004 cm^{-1}. The high-adipose-content cryosection (Figure 6.4B) has strong bands attributable to lipids, at 1745 and 1445 cm^{-1}, and proteins, at 1655 and 1300–1245 cm^{-1}, compared with the low-adipose-content cryosection (Figure 6.4C). The carotenoid bands are relatively weaker in the high-adipose-content cryosection than in the spectrum of the cryosection of low adipose content.

The FFPE sectioning technique uses over eight different chemicals during tissue processing [32]. The IR spectra from such sections demonstrated the importance of deparaffinisation prior to spectroscopic analysis.

Prior to FFPE sectioning, the tissue was embedded as a paraffin block. The Raman spectrum of the paraffin block (Figure 6.4G) is dominated by the peaks due to the long hydrocarbon chain, mainly from CH_2 modes. There are six characteristic bands; the two peaks at 1463 and 1441 cm^{-1} are assigned to the $\delta(CH_2)$ modes, whereas the band at 1418 cm^{-1} is assigned to the $\delta(CH_3)$ mode [33,34]. The band at ~1296 cm^{-1} is attributable to the $\delta(CH_2)$ mode, while the 1133 and 1063 cm^{-1} bands are due to the $\nu(C–C)$ modes [33,34].

Figure 6.4. Spectral evaluation of a pre-identified (histopathology) IDC lesions as a function of sample processing: (A) "bulk" tissue in PBS/D_2O; (B) cryosection containing adipose tissue; (C) cryosection without adipose tissue; (D) FFPE tissue block; (E) FFPE section; and (F) deparaffinised FFPE section. Raman spectra of chemical solutions (pathology grades) used in preparation of the FFPE sections: (G) paraffin wax; (H) xylene; and (I) histopure. Experimental conditions: (A–G) λ = 514-nm, ~2 mW at sample, exposure time between 10-60 s with 3 accumulations. (H and I) λ = 1064-nm; 200 scans; and 4 cm^{-1}. Bands: blue line: biological components; green line: β-carotene; grey dotted line: paraffin; and orange dotted line: deparaffinising agents. Spectra are baseline corrected and are offset for clarity.

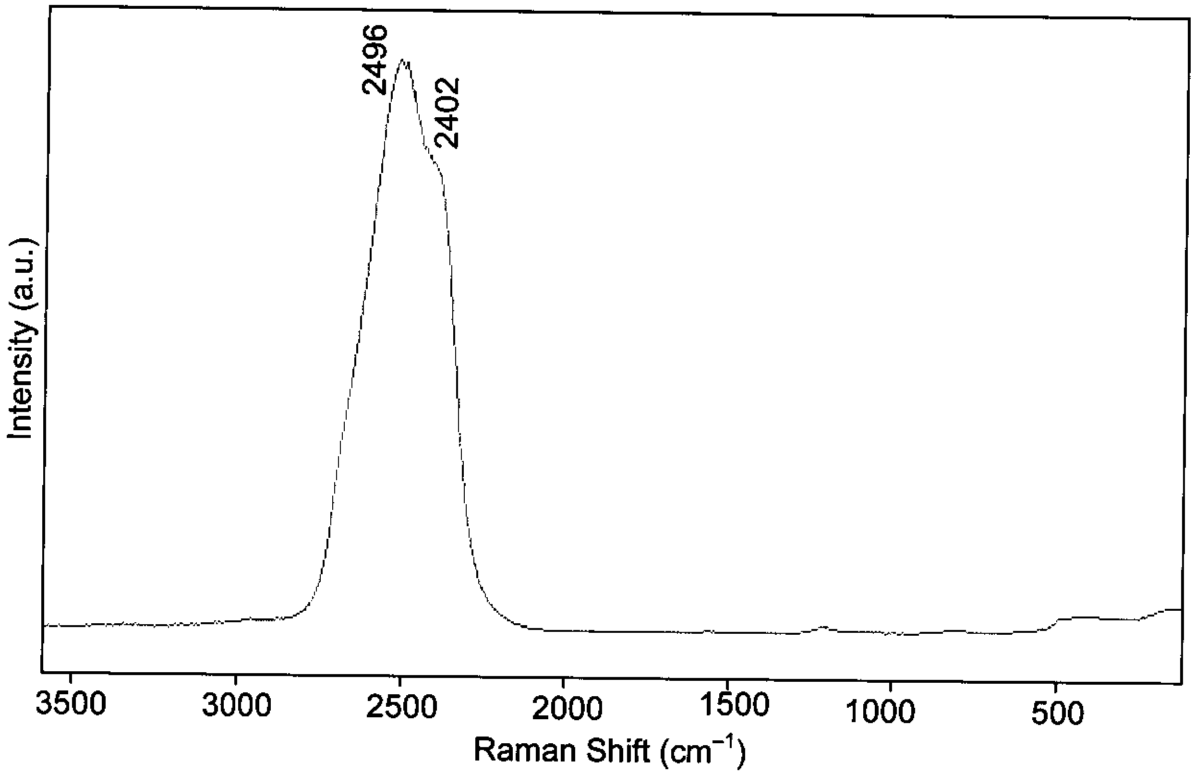


Figure 6.5. FT-Raman spectrum of PBS/D_2O solution showing two strong ν(O–D) bands at 2496 and 2402 cm^{-1}. Experimental conditions: 200 scans; 200 mW; and 4 cm^{-1}.

The spectrum measured from the tissue embedded in a paraffin block in Figure 6.4D exhibits strong interference from bands that are ascribable to paraffin. Visual examination of the tissue block did not reveal the presence of residual paraffin, but its presence is due to extensive infiltration during the overnight tissue processing (Table 2.1) prior to sectioning [32-34]. The paraffin is, therefore, retained within the tissue and the subsequent FFPE section prior to deparaffinisation (Figure 6.4E).

The FFPE section was then deparaffinised using xylene or histopure. The spectra of these two deparaffinising agents are presented in Figures 6.4H and I. After deparaffinisation using two baths of xylene and histopure for 10 min each, the paraffin removal was incomplete as shown by the paraffin bands in the spectrum from the 'deparaffinised' FFPE section (Figure 6.4G). The signature bands due to the paraffin are still visible (grey dotted lines), which is consistent with the IR spectroscopic findings (Section 4.4) and similar observations reported in the literature [33,34].

In addition, two bands at 1610 and 1630 cm^{-1} (orange dotted lines) are observed in the deparaffinised section, which are ascribable to the xylene solution (Figure 6.4H). The bands attributable to the biological components are observed after deparaffinisation but the spectra are inferior in quality to those obtained from the bulk and cryosectioned samples, due to chemical contaminations during FFPE tissue processing. As a result, deparaffinisation of FFPE is not an ideal sampling technique either for IR or Raman spectroscopy. Either fresh tissue or cryosections are the preferred samples for analytical techniques involving Raman spectroscopy.

6.1.4 MACROSPECTROSCOPIC APPROACH: TISSUE DEHYDRATION

In the last section, it was shown that both fresh tissue and cryosections were suitable for Raman spectroscopy. Macrospectroscopic approaches place the tissue biopsy, fresh or formalin-fixed, into a sample holder, which is illuminated by a laser (514, 830 or 1064 nm) at a specific laser power (~2 mW at sample with spot size of ~1 μm for 514 and 830 nm, and ~200 mW at the sample with laser spot size of ~100 μm for 1064 nm). The high laser power in a small sample area may alter the tissue structure over an extended period of laser exposure. In order to address this issue, a time-course study of a DCIS breast biopsy (~0.5 mm^3) exposed to a laser for 5, 10, 20, 40, 60 and 80 min was undertaken to investigate the effects of tissue dehydration using 1064-nm laser excitation (Figure 6.6). The sample was rehydrated with PBS/D_2O solution for 1 min in between each spectral collection.

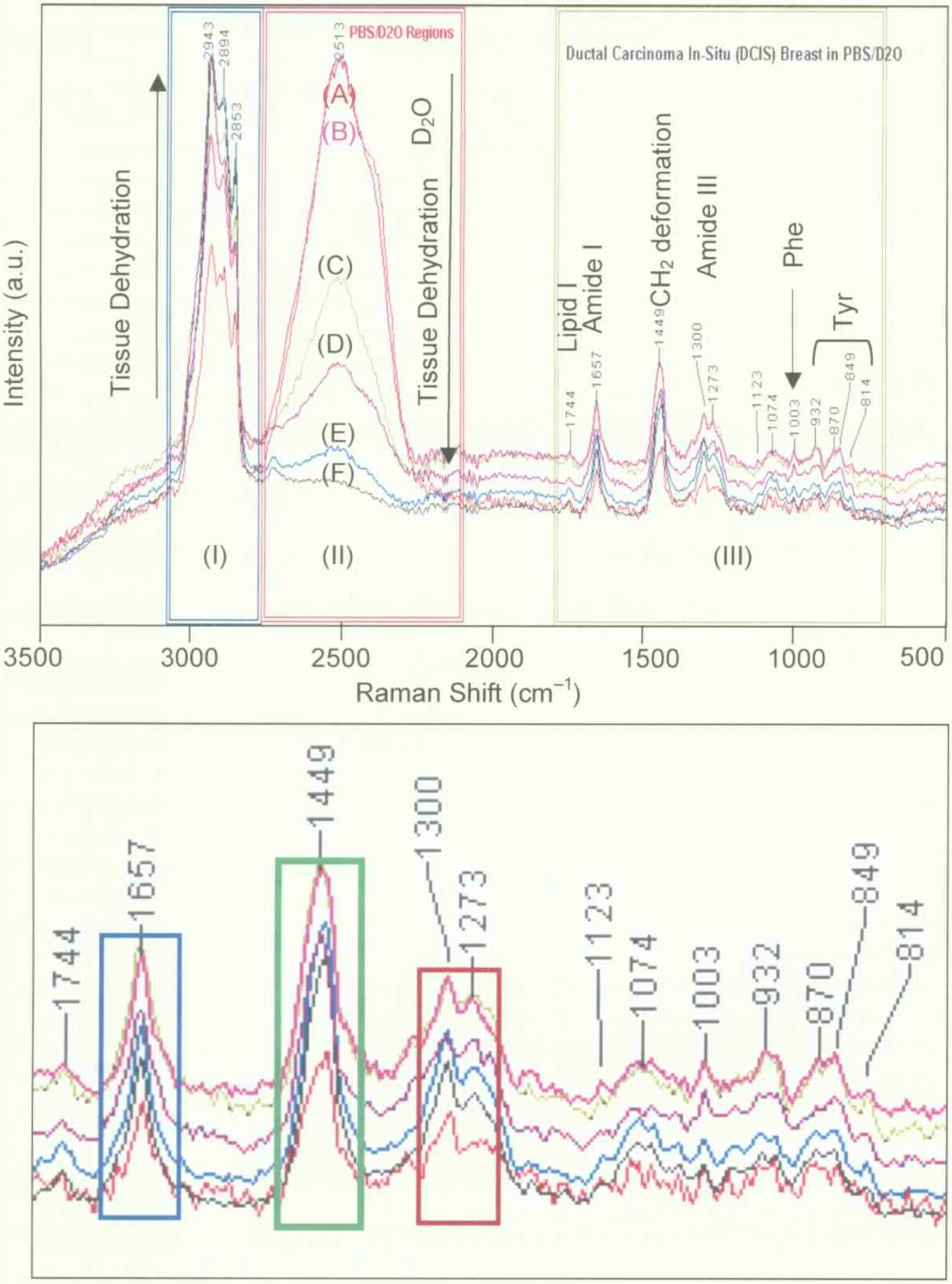


Figure 6.6. (Top) FT-Raman spectra at various exposure times of a DCIS breast biopsy stored in PBS/D_2O solution, then mounted on an aluminium sample holder: (A) 5 min (red); (B) 10 min (magenta); (C) 15 min (green); (D) 20 min (purple); (E) 30 min (blue); and (F) 60 min (black). (Bottom) Expanded 1750–750 cm^{-1}region. Experimental conditions: 200 mW; 256 scans; 4 cm^{-1}; and sample volume: 0.5 mm^3. Spectral variation observed in Region III is highlighted in boxes. The sample was discarded due to extensive tissue dehydration at the end of the experiment. Spectra were not baseline corrected.

Notable differences are observed in Regions (I) and (II). The intensity of the bands in the 3050–2750 cm^{-1} spectral region (Region I) increase relative to the water bands from PBS/D_2O, while there is a decrease in intensity of the ν(O–D) band of the D_2O solution in Region II over longer laser exposure times. A significant decrease in the ν(O–D) band is expected for extended laser exposure, while no significant differences are observed in the spectra from the samples exposed for 5 or 10 min (Figures 6.6A and B).

Although the biopsy dehydrated over time, the band intensities in Region III remain relatively unchanged. The expanded region (Figure 6.6 Bottom) shows that there is a slight change in the position and intensity of the amide III band at 1273 cm^{-1}. This change could be due to changes in protein structures over long exposure time as the tissue was dehydrated and such changes could be due to protein denaturation caused by sample heating. Although this would result in significant changes of hydrogen bonding involving the 'free water' and the protein (C=O···H–O–H), there was a slight change in the amide I band positions and the lineshapes in the spectra obtained at longer exposure times (20, 30, and 60 min). The lipid δ(CH_2) band shifts to high wavenumbers from the spectrum obtained at 5 min to the spectra obtained from longer exposure times, from 1447 to 1449 cm^{-1}.

After 40 min, the tissue had physically shrunk and hardened, and new DCIS biopsy samples with similar sample volumes (~0.5 mm^3) were used for each subsequent experiment. It was concluded that the laser exposure time using a macrospectroscopic approach should not exceed 15 min (equivalent to 512 scans at 200 mW).

6.1.5 MICROSPECTROSCOPIC APPROACH: TISSUE THICKNESS AND MOUNTING SUBSTRATE

For routine diagnosis using histopathology, IR and Raman spectroscopies, minimal variation in sample presentation is desirable. In IR spectroscopy, a standard 5- to 10-µm thickness resulted in a spectrum with good S/N ratios that could be used to differentiate breast lesions. Hence, it was necessary to examine whether the same tissue thickness could produce spectra of comparable quality in Raman spectroscopy. Raman spectra of a fibrocystic breast biopsy cryosectioned at thicknesses of 10, 20, and 50 µm and then mounted on CaF_2 windows, together with a bulk tissue and a standard 5-µm-thick cryosection mounted on transflective IR slide were compared using 830-nm laser excitation (Figure 6.7).

Clearly, no signal was observed for the 5-µm-thick cryosection mounted on a transflective IR slide (Figure 6.7E). The transflective slide reflects infrared light over the 1800–800 cm^{-1} spectral region and, therefore, is not suitable for Raman spectroscopy although the slide has a Ag/SnO_2 reflective surface. However, it has been demonstrated that quality spectra can be obtained from 10-µm-thick sections mounted on the transflective slide [33]. Similarly, there were no recognisable Raman bands from a sample of the same thickness mounted on a CaF_2 window, which is indicative of insufficient sample scattering for a thickness of 10 µm (Figure 6.7B).

In contrast, the bulk tissue (Figure 6.7A) produced comparable spectral information to those obtained from both the 20- and 50-µm-thick cryosections, Figures 6.7 C and D, respectively. As discussed in Chapter 4, thicker samples (>20 µm) are not recommended for spectroscopic diagnosis due to the high likelihood of signals originating from overlapping tissue components (ducts, lobules, adipose or stromal tissues) at any given point. In addition, thicker samples are usually more difficult to prepare. In this instance, the 20-μm-thick sample is the preferred thickness for breast tissue. In addition, bulk tissue analysis allows specific areas of interest, e.g., visual abnormalities, to be targeted. This allows Raman spectroscopy to be explored for routine *in-vivo* or *in-vitro* diagnostics.

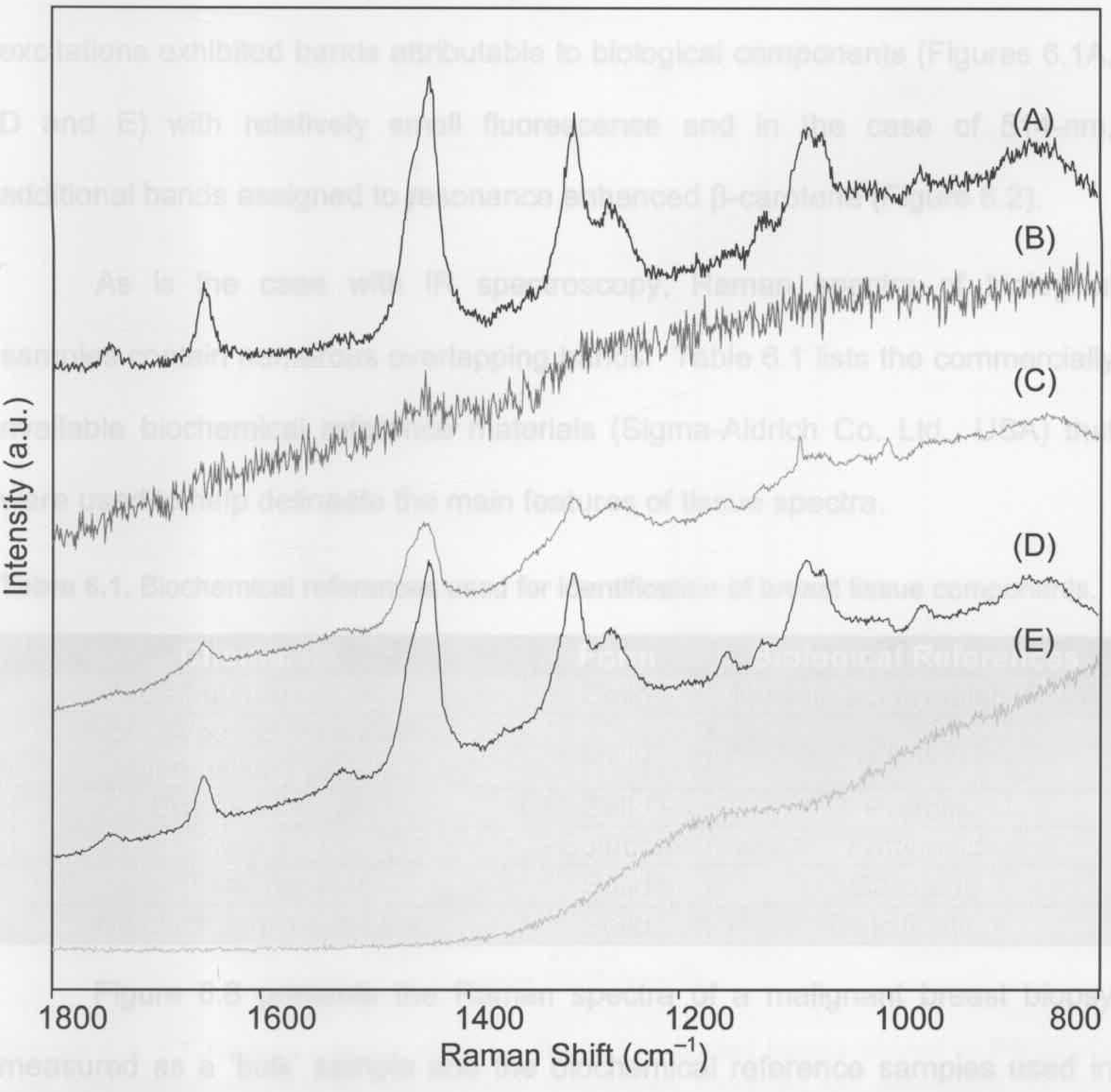


Figure 6.7. Raman spectral comparison of: (A) bulk fresh tissue measured at a suspicious black spot on the biopsy; cryosections mounted on CaF_2 window with various tissue thicknesses of: (B) 10-μm; (C) 20-μm; and (D) 50-μm; and (E) a standard 5-μm-thick cryosection mounted on transflective IR slide. Experimental conditions: λ = 830nm; ~2 mW at sample; 30 s; and 5 accumulations; and ×20 magnification. Spectra are offset for clarity.

6.2 SPECTRAL INTERPRETATION OF BREAST TISSUE USING FT-RAMAN SPECTROSCOPY

Raman spectra of breast specimens were collected with various laser excitations including: 514-, 647-, 752-, 830- and 1064-nm laser excitations. Section 6.1.1 described results that demonstrated that spectra collected using 647- and 752-nm laser excitations (Figures 6.1 B and C) were affected by a strong fluorescence background, which masked the biological bands of interest. On the other hand, the spectra acquired with 514-, 830-, and 1064-nm laser excitations exhibited bands attributable to biological components (Figures 6.1A, D and E) with relatively small fluorescence and in the case of 514-nm, additional bands assigned to resonance enhanced β-carotene (Figure 6.2).

As is the case with IR spectroscopy, Raman spectra of biological samples contain numerous overlapping bands. Table 6.1 lists the commercially available biochemical reference materials (Sigma-Aldrich Co. Ltd., USA) that were used to help delineate the main features of tissue spectra.

Table 6.1. Biochemical references used for identification of breast tissue components.

Chemicals	Form	Biological References
Calf thymus DNA	Solid	Nucleic acid (mainly DNA)
Mucin Type I-S from bovine submaxillary glands	Solid	Protein
Bovine albumin	Solid	Protein
Oleic acid	Solution	Fatty acid
β-carotene	Solid	Carotene
Calcium hydroxyapatite	Solid	Calcification

Figure 6.8 presents the Raman spectra of a malignant breast biopsy measured as a 'bulk' sample and the biochemical reference samples used in this research. The Raman spectrum will be discussed in three regions: (I) 4000–3050 cm^{-1}; (II) 3050–2750 cm^{-1}; and (III) 1800–500 cm^{-1}. Similar to IR spectroscopy, the 2750–1800 cm^{-1} region contains no significant diagnostic

information, i.e., biomolecules do not usually have bands in this region. The major bands and their assignments are summarised in Table 6.2. Depending on the laser excitations, certain bands were enhanced at particular laser excitations (e.g., carotenoids at 514-nm laser excitation). The band assignments are applicable to all laser excitations.

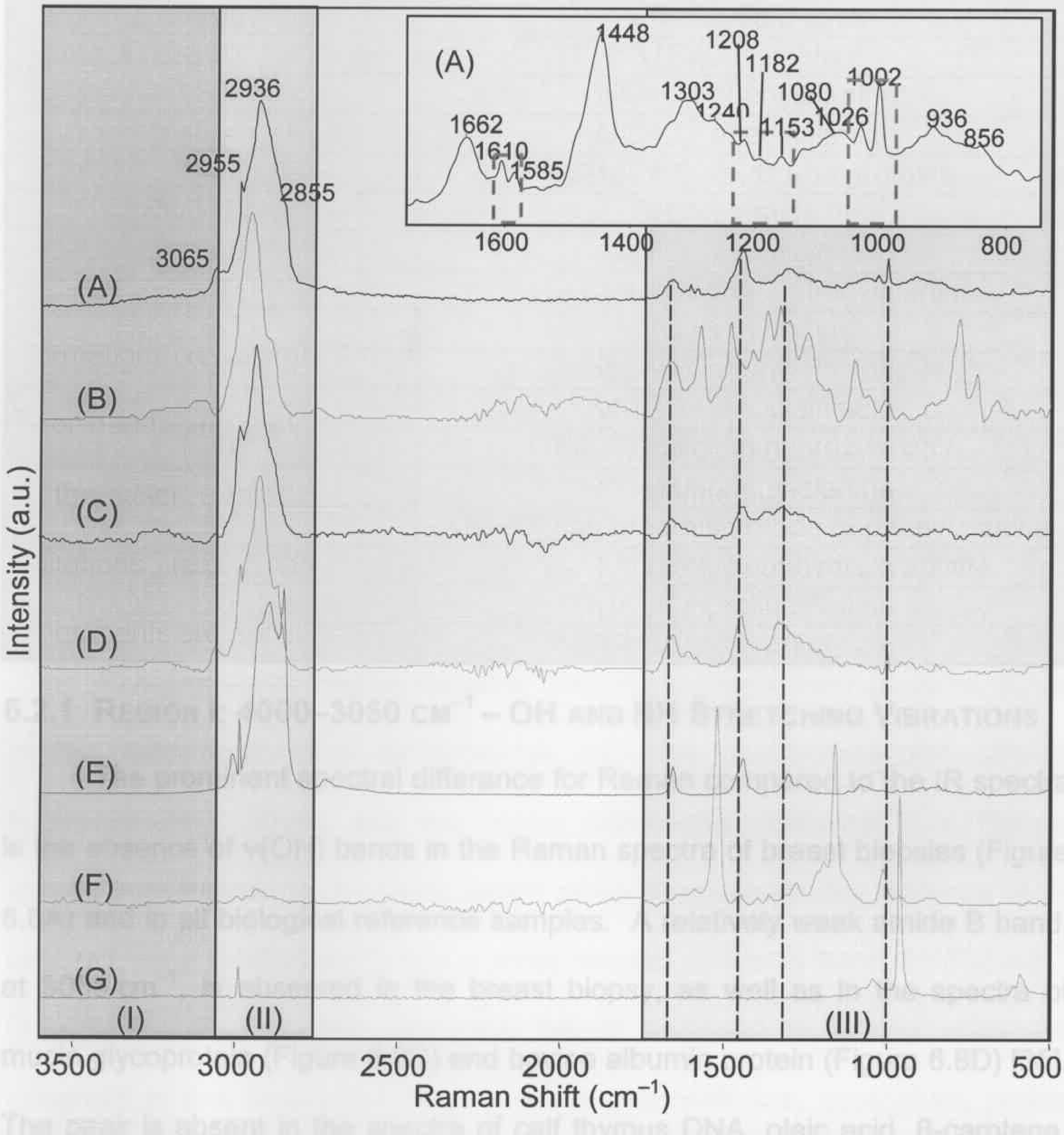


Figure 6.8. FT-Raman spectra of: (A) a fresh IDC breast biopsy stored in PBS/D_2O; and commercially-available biochemical references mounted on aluminium sample holders: (B) calf thymus DNA; (C) mucin type I-s glycoprotein; (D) bovine albumin protein; (E) oleic acid; (F) β-carotene and (G) calcium hydroxyapatite. Experimental Conditions: 200 mW; 200 scans; 4 cm^{-1}; and average of triplicates. (Inset) expanded 1800–800 cm^{-1}region of (A) an IDC breast biopsy. Red boxes are highlighting phenylalanine vibrational modes. Spectra were baseline-corrected and are offset for clarity.

Table 6.2. Major Raman peak assignments for breast tissues [35-38].

Wavelength (cm^{-1})	Peak Assignments
3300	ν(N–H), amide A
3065	δ(CNH), amide B
2955	$\nu_{as}(CH_3)$
2936	$\nu_{as}(CH_2)$
2896	$\nu_s(CH_3)$
2855	$\nu_s(CH_2)$
1735	ν(C=O), lipid ester
1662	ν(C=O), amide I
1607	δ(C=C), phenylalanine
1585	δ(C=C), phenylalanine
1448	$\delta(CH_2/CH_3)$, lipids and proteins
1300–1230	ν(CN), amide III
1240	$\nu_{as}(PO_2^-)$, nucleic acid
1208	ν(C–C), tryptophan, phenylalanine
1182	δ(C–H), tyrosine
1153	ν(C–N), phenylalanine
1080	$\nu_{as}(PO_2^-)$, nucleic acid
1044	$\nu(PO_4^{3-})$, calcium hydroxyapatite
1026	δ(C–H), phenylalanine
1002	ν_s(C–C) ring breathing mode of phenylalanine
960	$\nu(PO_4^{3-})$ from calcium hydroxyapatite
936	ν(C–C), proline, valine
856	δ(CCH) proline and tyrosine

6.2.1 Region I: 4000–3050 cm^{-1} – OH and NH Stretching Vibrations

The prominent spectral difference for Raman compared to the IR spectra is the absence of ν(OH) bands in the Raman spectra of breast biopsies (Figure 6.8A) and in all biological reference samples. A relatively weak amide B band, at 3065 cm^{-1}, is observed in the breast biopsy, as well as in the spectra of mucin glycoprotein (Figure 6.8C) and bovine albumin protein (Figure 6.8D) [35]. The peak is absent in the spectra of calf thymus DNA, oleic acid, β-carotene, and calcium hydroxyapatite (Figures 6.8B, E–G). The limited information in this region, therefore, is not useful for Raman diagnostics.

6.2.2 REGION II: 3050–2750 CM^{-1} – CH STRETCHING VIBRATIONS

The 3050–2750 cm^{-1} spectral region contains peaks that are assigned to the various CH symmetric and anti-symmetric stretching vibrations [35]. Similar to the IR spectrum, these bands are due to the combination of proteins, lipids, nucleic acids and chemicals used in tissue processing.

Based on the literature, the peaks at 2955 and 2896 cm^{-1} are assigned to the $\nu_{as}(CH_3)$ and $\nu_s(CH_3)$ modes whereas those of 2936 and 2855 cm^{-1} are assigned to $\nu_{as}(CH_2)$ and $\nu_s(CH_2)$ modes, respectively [35,39]. The relative intensity and line-shape of this spectral region is highly dependent upon the amount of the individual tissue components present in the sample. The intensities of $\nu_{as}(CH_2)$ and $\nu_s(CH_2)$ bands are elevated in the spectra of tissue with high adipose content, similar to those in the oleic acid spectrum (Figure 6.8E).

6.2.3 REGION III: 1800–500 CM^{-1} – FINGERPRINT REGION

The fingerprint region contains diagnostic bands from proteins, lipids, nucleic acids and other biomolecules including carotenoids and calcium hydroxyapatite. There are two major differences observed in a Raman spectrum compared to the same region of an IR spectrum: the absence of an amide II band (~1550 cm^{-1}) and a more intense amide III band (~1330 cm^{-1}) [37,40-42]. Furthermore, additional bands assigned to amino acids are observed in this region including: tyrosine, phenylalanine, tryptophan, proline and valine.

The band at 1662 cm^{-1} is assigned to the amide I mode originating mainly from proteins and nucleic acids (Figures 6.8B–D). The two weak bands at 1610 and 1585 cm^{-1} observed in the breast tissue (red box) and bovine albumin protein (Figures 6.7A and D) are due to the δ(C=C) modes of aromatic amino acids (phenylalanine, tyrosine, and tryptophan) [37,38,43]. The band at 1448 cm^{-1} is assigned to the δ(CH_2/CH_3) modes from a combination of lipoproteins from the cell membrane, adipose tissue, and nucleic acids. The amide III bands are observed in the region of 1295–1200 cm^{-1}, which are attributed to a combination of ν(CN) and δ(NH) modes of the peptide bond (–CONH). Other bands assigned to the amino acids include: the ν(C–C) modes of tryptophan and phenylalanine at 1208 cm^{-1}, the δ(C–H) mode of tyrosine at 1182 cm^{-1}, the ν(C–N), δ(C–H) and ν_s(C–C) ring breathing modes of phenylalanine at 1153, 1026 and 1002 cm^{-1}, respectively [37,38,43]. The bands at 936 and 856 cm^{-1} are assigned to the ν(C–C) modes of proline and valine, and the δ(CCH) modes of proline and tyrosine, respectively [37,38,43].

The spectrum of β-carotene (Figure 6.8F) exhibits three major characteristic bands in this region including those due to: the ν(C=C) mode at 1515 cm^{-1}, the ν(C–C) mode at 1156 cm^{-1}, and the ring breathing mode at 1004 cm^{-1} [12]. Although there is a minimal contribution from β-carotene in the spectrum of the biopsy measured with 1064-nm laser excitation, these bands are resonance-enhanced using 514-nm laser excitation (Figure 6.2).

Two major bands are observed in the calcium hydroxyapatite ($Ca_{10}(OH)_2(PO_4)_6$) spectrum (Figure 6.8G) at 1041 and 961 cm^{-1}, which are assigned to the $\nu_{as}(PO_4^-)$ and $\nu_s(PO_4^-)$ modes, respectively [44].

6.3 DIFFERENTIATION OF BREAST LESIONS

Differentiation of breast lesions is based on macrospectroscopic approaches using FT-Raman spectroscopy or microspectroscopic approaches using dispersive Raman microspectroscopy with laser excitations of 514 and 830 nm. Raman spectra using a macrospectroscopic approach were collected by placing the biopsy samples on aluminium sample holders, whereas 20-μm-thick tissue sections were mounted on either CaF_2 windows or mirrors with the microspectroscopic approach. Due to the limited availability of the 830-nm laser (on loan from The Medical Laser Institute, The University of Liverpool, UK), only two IDC samples were analysed, with the focus being identification of micro-calcification using Raman imaging.

6.3.1 BULK TISSUE ANALYSIS USING MACROSPECTROSCOPIC APPROACH

Representative Raman spectra from a benign fibrocystic change and a malignant IDC biopsy fixed in formalin are presented in Figure 6.9. These spectra show that a 10-minute (or equivalent to 512 co-added scans at 4 cm^{-1} resolution) collection time produces good S/N ratios without significant sample dehydration. However, this approach was not used in subsequent experiments as the spatial resolution was inappropriate. The spot size for the FT-Raman operating in macroscopic and microscopic modes are approximately 100 and 65 μm, respectively. Furthermore, the small sample size (~1 mm^3) made sectioning for further histopathological and spectroscopic (IR and Raman) analyses difficult.

6.3.1.1 REGION I: 4000–3050 CM^{-1}

The amide B band at 3065 cm^{-1} is observed to be stronger relative to other bands in that of the spectrum of the benign fibrocystic change lesion (Figure 6.9A) than in the malignant IDC lesion (Figure 6.9B). However, this region lacks diagnostic spectral information (absence of amide and H-bonding water bands) for tissue comparison and is subsequently not used for diagnosis.

6.3.1.2 REGION II: 3050–2750 CM^{-1}

The positions of the ν(CH) stretching vibrations for both benign and malignant biopsies are similar. The major difference is the intensities of the $\nu_s(CH_3)$ and $\nu(CH_2)$ bands at 2896 and 2855 cm^{-1} in the biopsy from a malignant lesion. The higher intensities of these two bands relative to others are indicative of a higher lipoprotein content due to cellular proliferation with a small contribution of adipose tissue in the formalin-fixed IDC biopsy.

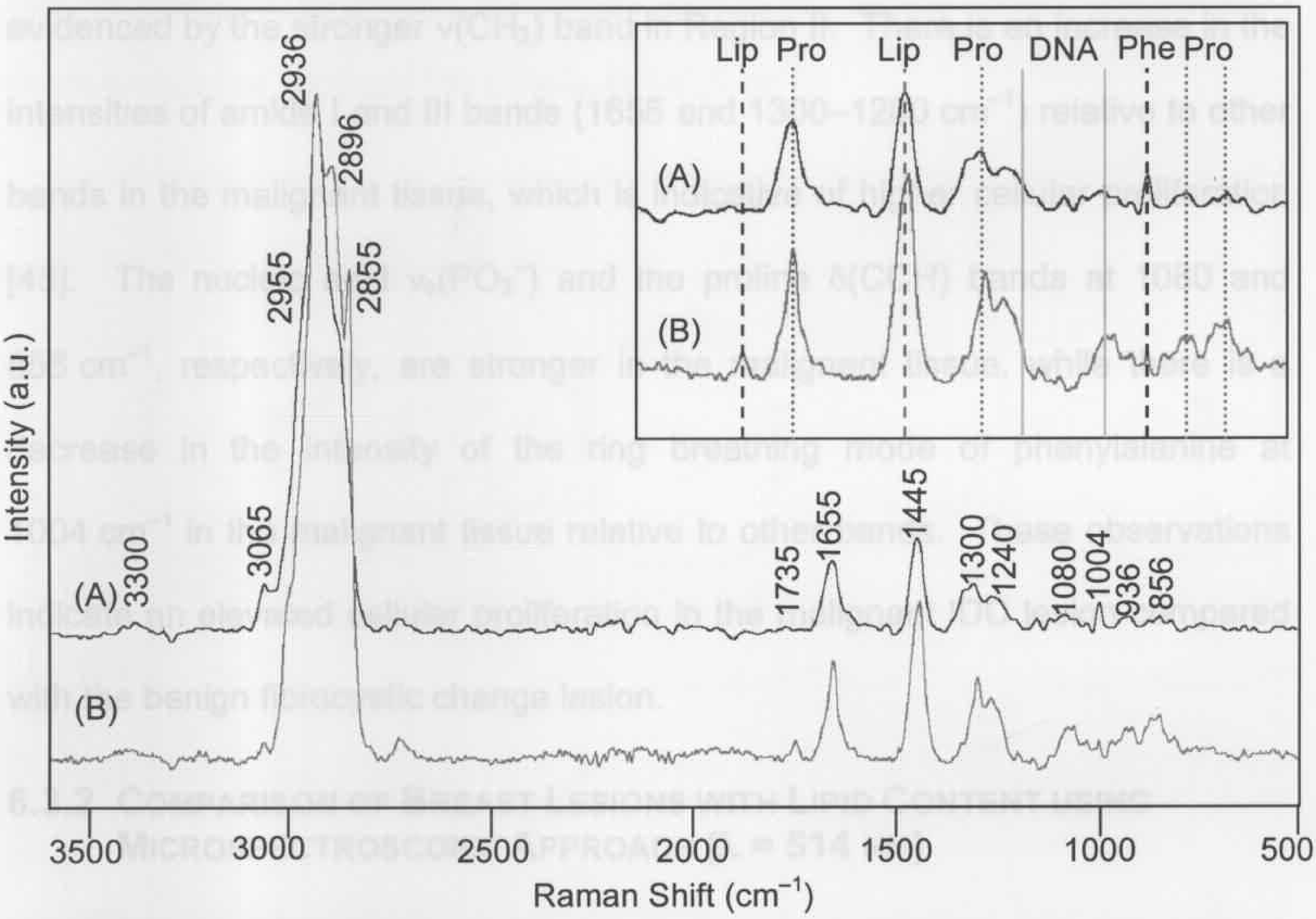


Figure 6.9. Raman spectra of formalin-fixed human breast biopsies: (A) benign fibrocystic change lesion and (B) malignant IDC lesion. (Inset) Expansion of 1800–850 cm^{-1} region showing the major biological components. Lip: lipid; Pro: protein; DNA: deoxyribose nucleic acid; and Phe: phenylalanine. Experimental conditions: λ = 1064-nm; 200 mW, 512 scans; and 4 cm^{-1}. Spectra are offset for clarity.

6.3.1.3 REGION III: 1800–500 CM^{-1}

As discussed in Chapter 4, formalin fixation retains the adipose content but the amide I band shifts to a lower wavenumber due to the alteration of protein secondary structure during cross-linking. However, the amide I band at

1655 cm^{-1} was not observed to shift to lower wavenumbers in the Raman spectra of either the benign or malignant biopsies compared with the spectra of DCIS stored in PBS/D_2O (Figure 6.6) These findings are consistent with the previous work by Frank *et al.* [12].

The major differences in this region between the malignant tissue spectra are the presence of a lipid ester $\nu(C=O)$ band at 1735 cm^{-1}, and more intense $\delta(CH_2)$ bands at 1445 cm^{-1}, which are due to the adipose tissue as evidenced by the stronger $\nu(CH_2)$ band in Region II. There is an increase in the intensities of amide I and III bands (1655 and 1300–1200 cm^{-1}) relative to other bands in the malignant tissue, which is indicative of higher cellular proliferation [45]. The nucleic acid $\nu_s(PO_2^-)$ and the proline $\delta(CCH)$ bands at 1080 and 856 cm^{-1}, respectively, are stronger in the malignant tissue, while there is a decrease in the intensity of the ring breathing mode of phenylalanine at 1004 cm^{-1} in the malignant tissue relative to other bands. These observations indicate an elevated cellular proliferation in the malignant IDC lesion compared with the benign fibrocystic change lesion.

6.3.2 COMPARISON OF BREAST LESIONS WITH LIPID CONTENT USING MICROSPECTROSCOPIC APPROACH (λ = 514 NM)

Raman spectra of four 20-μm-thick breast cryosections (two benign fibrocystic changes and two malignant IDC) mounted on CaF_2 windows and a reference β-carotene were collected using 514-nm laser excitation (Figure 6.10). The spectra were collected from the tissue surrounding the duct and are normalised to the strong intensity of the carotenoid $\delta(CH_2)$ band at 1514 cm^{-1}.

The resonance enhanced carotene bands at 1514, 1155 and 1004 cm^{-1} dominate all cryosection spectra. The intensities of the amide I and III bands, at

1655 and 1300 cm^{-1}, and lipid bands, at 1740 and 1445 cm^{-1}, are relatively stronger in the spectra of malignant cryosections compared to the benign cryosections. In the high-lipid malignant cryosection (Figure 6.10A), the amide I band is centred at 1655 cm^{-1}, while the band shifts to lower wavenumber (~1638 cm^{-1}) in the spectra of both benign cryosections (Figures 6.10B and D). The low-lipid malignant cryosection (Figure 6.10C) exhibits a weak band at 1655 cm^{-1} but a relatively stronger band at 1638 cm^{-1}. The changes could be due to changes in protein secondary structures during cellular proliferation (from β-sheet at 1638 cm^{-1} to α-helix at 1655 cm^{-1}).

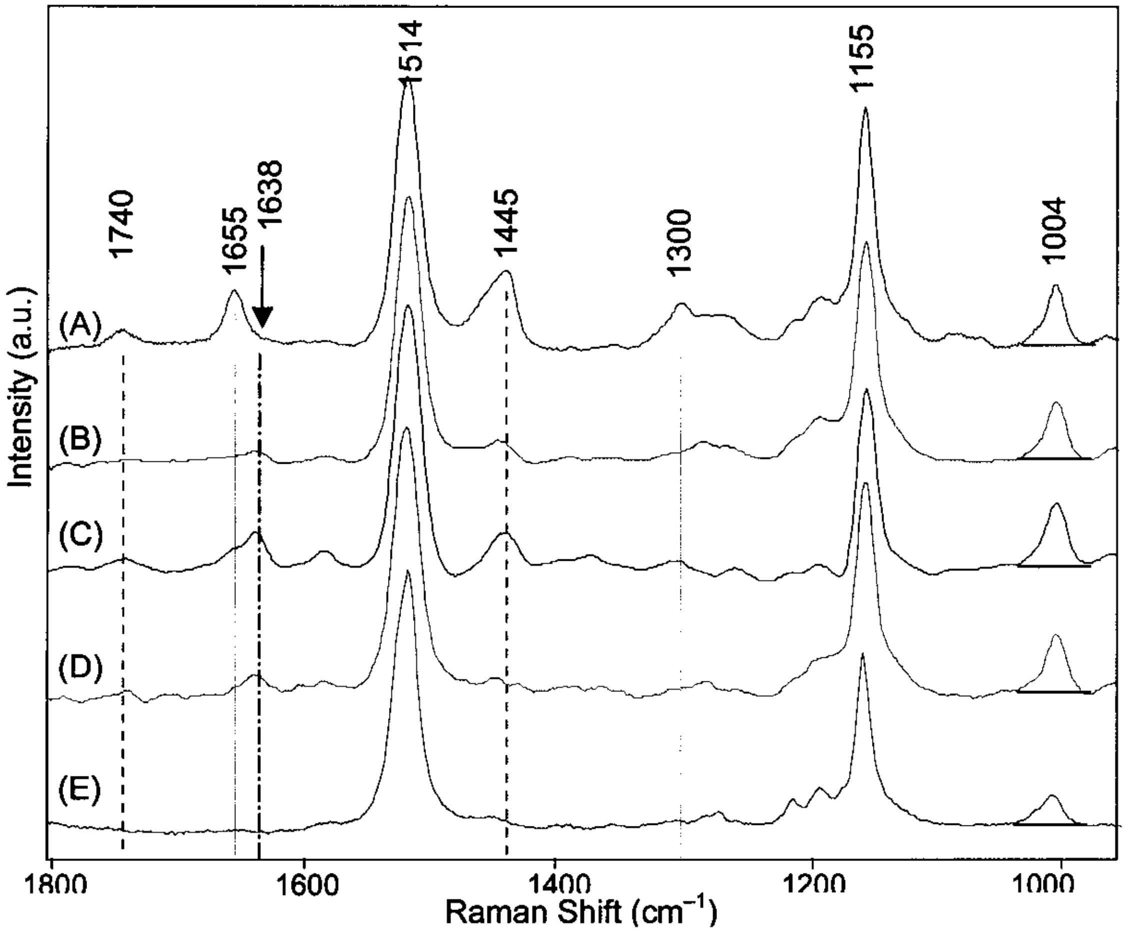


Figure 6.10. Microspectroscopic comparison of lipoprotein levels in: (A) a high-lipid malignant IDC cryosection; (B) a high-lipid benign fibrocystic change cryosection; (C) a low-lipid malignant IDC cryosection; (D) a low-lipid benign fibrocystic cryosection; and (E) β-carotene. Experimental conditions: λ = 514 nm; (A–D) 2 mW at samples; 30 s; 1 accumulation; ×20 magnification; and thicknesses of 20 μm; and (E) 0.2 mW; 10 s; 1 accumulation; and ×20 magnification. Spectra are normalised to the β-carotene band at 1514 cm^{-1}, baseline corrected and are offset for clarity.

The amide III (1300 cm^{-1}) and lipid (1740 and 1445 cm^{-1}) bands are relatively stronger in both malignant cryosections than in the spectra of benign cryosections. The stronger intensities of lipid in both malignant cryosections are indicative of an increase in lipoprotein levels from the cell membrane. The major pathological feature of the benign fibrocystic change is an enlargement of the duct due to the cyst formation. No cellular proliferation is present in this condition and the major lipid bands are mainly due to the adipose tissue [46].

6.4 CHARACTERISATION AND IDENTIFICATION OF LCIS AND MICRO-CALCIFICATION USING RAMAN IMAGING

Raman imaging uses filters to acquire the image, which bypass the grating used in single-point microspectroscopy. The laser spot size is expanded to various defocus percentages during the experiment and is controlled by the computer software [47]. The higher defocus percentages result in a larger sampling area for imaging. Prior to imaging, a filter spectrum was collected to locate the band of interest, e.g., the amide I band at 1650 cm^{-1}, which may shift ± 10 cm^{-1} when acquired using the grating, and the strongest peak from the filter spectrum is then used to collect a Raman image. The acquisition time in this study was 600 s or ten times the collection time used to acquire single-point grating spectrum. Two laser excitations, 514 and 830 nm, were employed: 514-nm excitation for optimisation and characterisation of a LCIS cryosection, while the 830-nm excitation was employed for identification of micro-calcification in a malignant IDC cryosection.

6.4.1 OPTIMISATION FOR RAMAN IMAGING: β-CAROTENE

The β-carotene bands exhibit strong resonance Raman enhancement on excitation of the $\pi \rightarrow \pi^*$ electronic transition using 514-nm laser excitation [17].

Prior to an investigation of breast lesions, β-carotene samples were imaged to establish the optimal conditions for subsequent Raman imaging. The photomicrograph and a representative Raman spectrum of β-carotene deposited on an aluminium sample holder are presented in (Figures 6.11A and B). The Raman spectrum exhibits strong resonance-enhanced bands at 1514 and 1155 cm^{-1} ascribed to the ν(C=C) and ν(C–C) vibrations, respectively, and the ν(C=C) band was used to obtain Raman images (Figure 6.11B) [12,17].

Ten Raman filter images centred at 1508 cm^{-1} (based on the 1514 cm^{-1} ν(C=C) band from the grating spectrum) with various defocusing percentages (0 to 50%) using 514-nm laser excitation are depicted in Figure 6.11C. The images all have a calculated spatial resolution of 1.6 μm for the ×20/0.40 NA objective lens at 514-nm laser excitation. Regions with high intensities of β-carotene distribution are in red, while regions without β-carotene or where the laser beam did not illuminate the sample are represented in black. The area of high intensity (red) in the 0% defocused image is due to uneven packing of the sample, which was unavoidable and a consequence of the granular nature of the sample. The β-carotene produced the best images when the laser was defocused between 5 and 15%. The samples underwent photo-degradation at defocusing percentages of 35 to 50% as evident by a change in colour from red-brown to black. However, the images with 20–25% defocusing show that the laser beam has illuminated to the maximum image area compared to the image acquired with defocusing of 0%. As a result of this preliminary experiment, Raman images on tissue analysis were acquired with the laser beam defocus set between 5 and 20%.

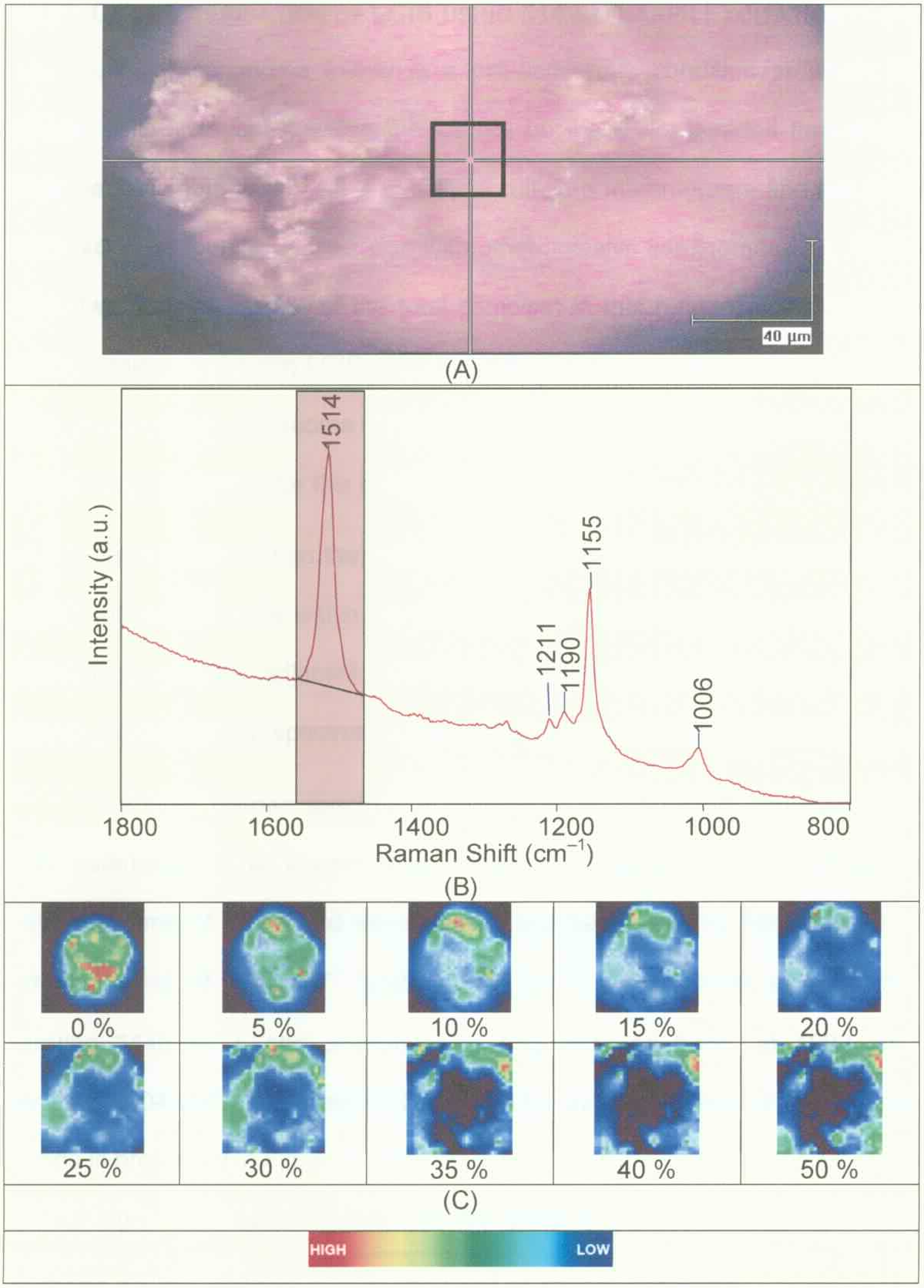


Figure 6.11. (A) Photomicrograph of β-carotene deposited on an aluminium sample holder and (B) the Raman spectrum collected using 514-nm laser excitation; and (C) *the Raman images using the* ν*(C=C) band at 1514 cm*$^{-1}$ *with various laser defocus* percentages. Experimental conditions: λ = 514-nm; laser power: 0.1 mW at sample; 600 s; 1 accumulation; and ×20 magnification. Spatial resolution: 1.6 μm and image dimension: 28 × 28 μm^2.

6.4.2 CHARACTERISATION OF LCIS USING 514-NM LASER EXCITATION

Lobular carcinoma *in-situ* is a pre-cancerous condition in which the tumour cells are located within the lobules but have not invaded through the basement membrane [46]. It is usually occult* in a mammogram and has been reported to be found in association with other invasive carcinomas [49]. There were two specimens (2% of the total samples) in this research diagnosed by histopathology as having LCIS. However, there is only limited research on this rare pre-invasive carcinoma [11]. Therefore, a better understanding is important for detection for this condition.

The box marked on the unstained section (Figure 6.12A Left) contains a cluster of tumour cells within the lobules, and was identified as LCIS by a pathologist using the adjacent H&E stained section (Figure 6.12A Right). The representative Raman spectrum of the LCIS is presented in Figure 6.12B.

The LCIS Raman spectrum is dominated by carotenoid bands (marked with asterisks). The images were collected using 20% laser defocus, a collection time of 600 s and were constructed based on the filter spectra of amide I band, at 1650 cm^{-1} (grating 1654 cm^{-1}), $\delta(CH_2)$ band, at 1435 cm^{-1} (grating 1440 cm^{-1}), and phenylalanine ring-breathing band, at 1002 cm^{-1} (grating 1004 cm^{-1}), (Figures 6.12D–F). The overlaid image of these three bands is presented in Figure 6.12C. The image colours were inverted for clarity and a higher distribution of a particular band is represented by light blue, whereas a lower distribution is represented by yellow/pale-green.

---

* Occult is the medical term for difficult to see, not visible to the naked eye, and only detectable by microscope or chemical testing [48] *Oxford Concise Medical Dictionary*. 6th ed., Oxford University Press, Oxford, 2003.

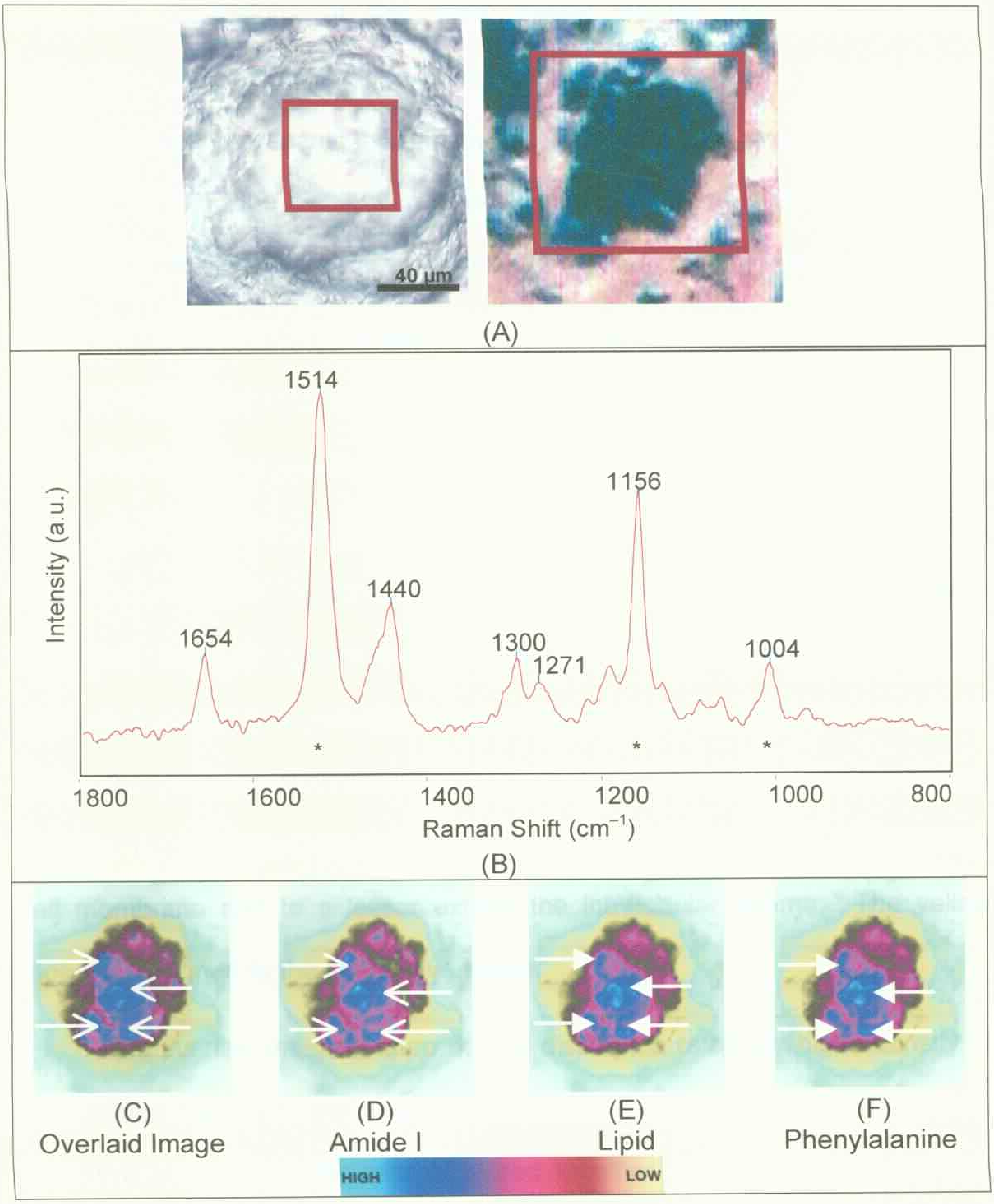


Figure 6.12. (A) Photomicrograph of a cluster of tumour cells within the lobules of an unstained 20-μm-thick LCIS cryosection mounted on CaF$_2$ window, and (right) the adjacent H&E stained section. The boxes indicate the areas where the Raman images were collected. (B) The representative LCIS Raman spectrum. Bands assigned to the β-carotene are marked with asterisks. Raman images constructed using: (C) the overlaid image of D–F; (D) amide I (1654 cm^{-1}) band; (E) lipid (1440 cm^{-1}) band; and (F) Phe (1004 cm^{-1}) band. Arrows: high level of tumour cells (nuclei). Experimental conditions: 514-nm laser excitation, filter image 600 s, 20% laser defocus. Spectrum was measured using 1800 lines/mm grating, 10s, 1 accumulation, 1800–800 cm^{-1} region, 20% laser defocus, x20 magnification. Spatial resolution: 1.6 μm and image dimension: 58 × 58 μm^2. The spectrum is baseline corrected and the image colours are inverted for clarity.

The dimensions of the small breast lobules were approximately 120 × 120 μm^2, and the images produced were 58 × 58 μm^2. Therefore, the image area only represents the tumour cells within the LCIS lobule. As seen in Figures 6.12C–F, the distribution of each diagnostic band is similar but varied within the image, which is due to variation of cellular components.

There are four major areas containing a high amide I distribution (arrows, Figure 6.12D), while the remaining areas show medium intensity (magenta). A similar distribution is observed in lipid and carotene images. These consistent areas of high concentration are indicative of nuclei within the tumour cells, which are clustered together, as shown in the H&E stained image (Figure 6.12A). In the area coloured in blue, there are relatively higher intensities of the amide I, lipid and phenylalanine bands within the tumour cells, which could be due to the presence of nuclei, as evidenced of the H&E stained section (Figure 6.12A). The medium intensity observed around the nuclei could be due to the cell membrane and to a lesser extent, the intralobular stroma. The yellow region is assigned to the connective tissue.

The overlaid image (Figure 6.12C) displays a relatively high intensity of these three bands in the areas indicated by arrows, and these areaa are indicative of more tumour cells clustered together compared to the top-right area. As shown here, Raman imaging could be used to characterise various tissue components for small lesions, such as lobular carcinoma *in-situ*.

6.4.3 IDENTIFICATION OF MICRO-CALCIFICATION IN AN IDC LESION

Location of masses and calcifications are the main characteristics of a lesion on a diagnostic mammogram. Masses can be due to a cancer or a benign condition, while micro-calcifications are small mineral deposits within the breast tissue that appear as small white spots on the mammograms [46,50,51]. There are two major types of micro-calcifications that have similar morphological characteristics when viewed on a mammogram, type I calcium oxalate dihydrate ($CaC_2O_4 \cdot 2H_2O$) and type II calcium hydroxyapatite ($Ca_{10}(OH)_2(PO_4)_6$) [46,52]. The type I micro-calcification is normally found in benign breast lesions and is associated with cellular secretions, whereas type II micro-calcification is associated with malignant breast cancer as a result of necrosis [46,52].

Raman spectroscopy has been used to identify micro-calcification in various cancers [1,2,11,53]. These results were based on the identification of signature bands from both types of micro-calcification using single-point spectroscopy. Koljenovic and co-workers applied k-means cluster analysis to image the type II calcification deposit in a necrotic brain tissue using the intense calcium hydroxyapatite band at 961 cm^{-1} [53]. However, there is limited research on the micro-calcification on breast cancer [2].

Raman spectra of an invasive ductal carcinoma with micro-calcification and a commercial sample of calcium hydroxyapatite are presented in Figure 6.13. A reference spectrum of calcium oxalate was previously published by Shafer-Peltier *et al.* and contains two major bands at 920 and 1480 cm^{-1} [2]. Two bands observed at 1044 and 961 cm^{-1} are attributed to the $\nu(PO_4^{2-})$ modes of calcium hydroxyapatite in the invasive ductal carcinoma breast lesion (Figure 6.13A), as shown by comparison with the Raman spectrum of calcium hydroxyapatite (Figure 6.13B) [2,53]. The calcification was also confirmed by the H&E stained section (Figure 3.12). This observation, therefore, can be used to determine the presence and the type of micro-calcification (type I or II) to diagnose the nature of a breast lesion condition.

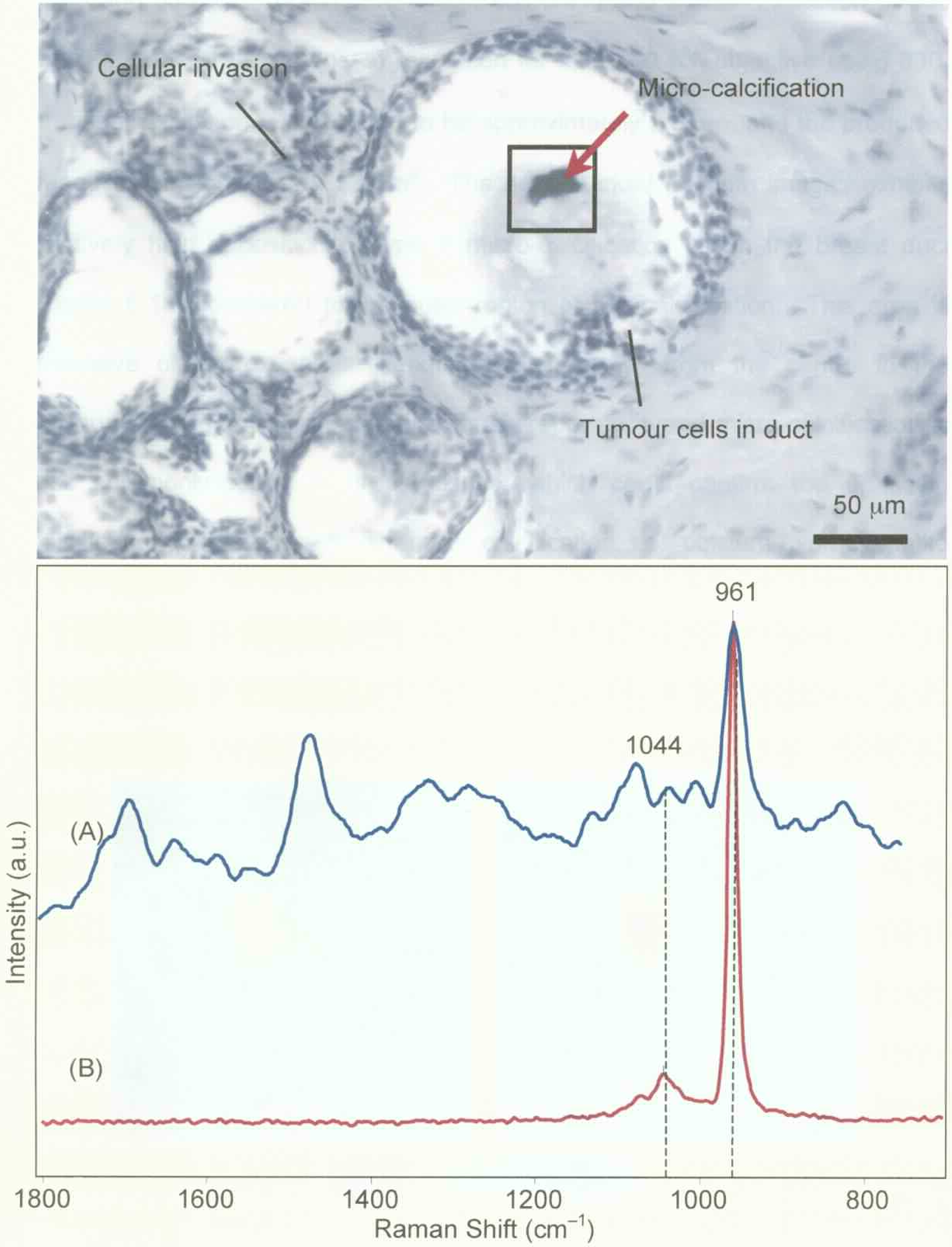


Figure 6.13. (Top) A H&E-stained section of an IDC with calcification (arrow) and the area (box) where Raman spectrum and images were collected. (Bottom) Raman spectra of: (A) an invasive ductal carcinoma breast lesion with micro-calcification; and (B) a reference spectrum of calcium hydroxyapatite. Experimental conditions: λ = 830 nm, laser power: ~2 mW at sample; average of triplicates. Spectra are baseline-corrected and are offset for clarity.

Two Raman images were constructed using the hydroxyapatite bands at 1040 and 958 cm^{-1} (grating: 1044 and 961 cm^{-1}) and a laser excitation of 830 nm (Figure 6.14). The spatial resolution for ×20/0.40 NA objective using 830-nm laser excitation is calculated to be approximately 2.5 μm, and the produced image dimension is 58 × 58 μm^2. The area marked on both images exhibits relatively high deposition of type II micro-calcification within the breast duct (Figure 6.13) compared to the outer region of the calcification. This area is indicative of hydroxyapatite seeding, which grows from the centre to the surrounding region (green). As previously reported, type I micro-calcification is less commonly found in breast tissue, which could confirm the unknown calcification is due to the type II micro-calcification, i.e., calcium hydroxyapatite [2]. This result demonstrates that Raman imaging is capable of detecting micro-calcifications and has the potential of differentiating benign from malignant breast lesions.

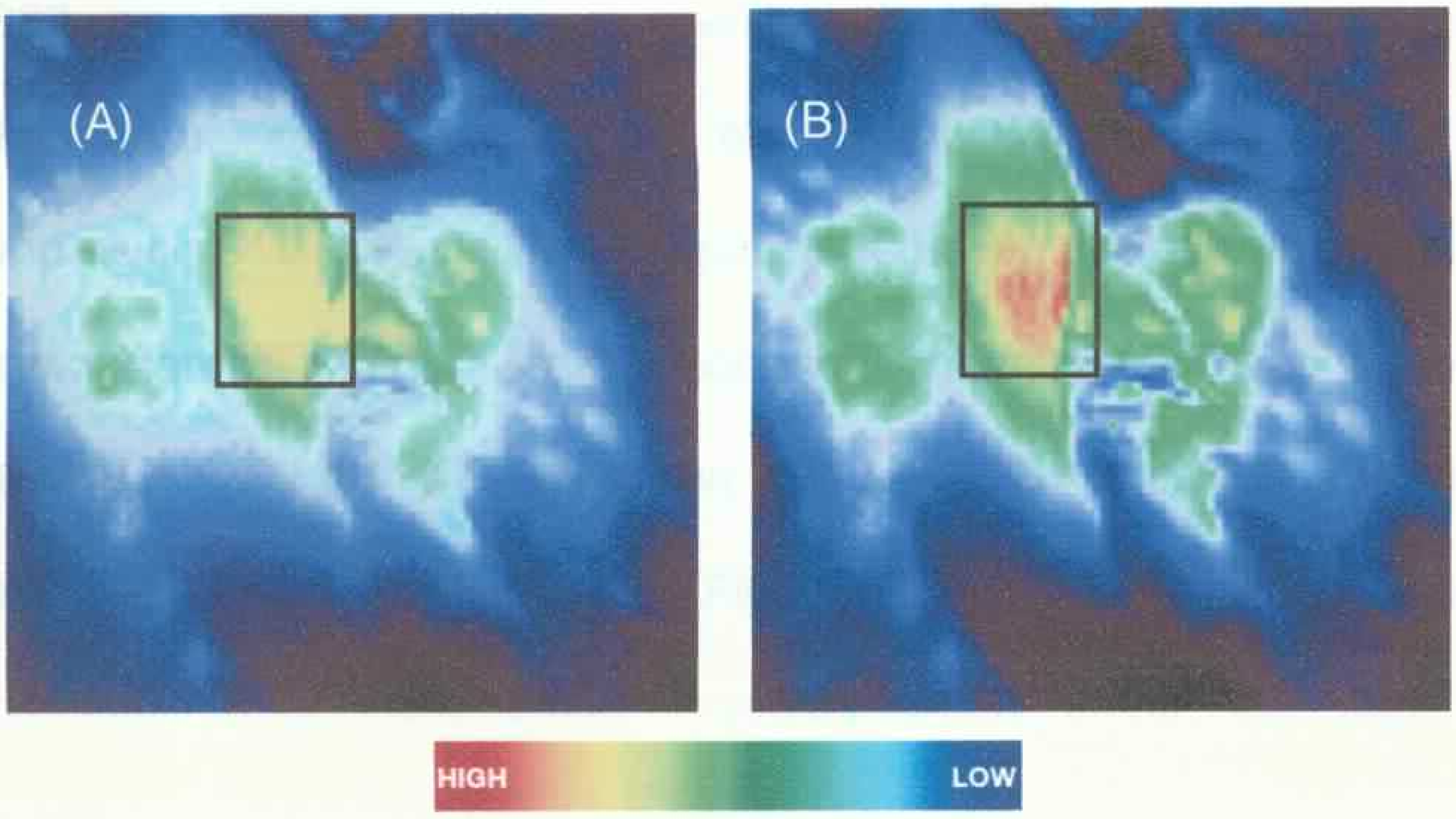


Figure 6.14. Raman images of an invasive ductal carcinoma breast cryosection in Figure 6.13 showing type II micro-calcification constructed using intensities of the bands at: (A) 1044 cm^{-1}; and (B) 961 cm^{-1}. The box is the area where there is high concentration in calcium hydroxyapatite of the micro-calcification. Experimental conditions: λ = 830-nm; 600 s; ×20 magnification; spatial resolution of 2.5 μm; and image dimension of 58 x 58 μm^2.

6.5 CONCLUSIONS

Raman spectroscopic investigations of breast tissues have been extensively reported mainly using near-infrared laser excitations [1,2,9,10,12,14,17,21,40,42,54]. This research had focussed on three major aspects: sample optimisation, macrospectroscopic and microspectroscopic differentiation of breast tissues using 514- and 830-nm laser excitations, and characterisation of the tumour-filled breast lobule of a pre-defined LCIS lesion and identification of micro-calcification in an IDC lesion using NIR-Raman imaging techniques.

Initial studies on the selection of appropriate laser excitation (λ = 514, 647, 752, 830 and 1064 nm) showed various degrees of spectral information could be obtained at different wavelengths. However, the strong fluorescence background generated using 647- and 752-nm excitation masked the Raman features of biological components. Subsequently, the research focussed on the 514-, 830-, and 1064-nm excitation wavelengths. Different spectral information was obtained using these three laser excitations. The main spectral features observed using 514-nm excitation (Figure 6.2C) were the resonance-enhanced carotenoid Raman bands. This resonance enhancement was weak (post-resonance) using a 830-nm excitation (Figure 6.2B) and was absent in the spectrum using a 1064-nm excitation (Figure 6.2A). However, the bands due to proteins and lipids were better resolved when using the near-infrared excitations, which could be used for diagnostic comparison.

Since water is a weak Raman scatterer, Raman spectroscopy can be used in pathology laboratories to monitor the physical state of tissue sections. As shown in Figure 6.4, spectral contamination arising mainly from paraffin wax was observed in the spectrum of deparaffinised FFPE section. This demonstrates that the recommended deparaffinising time in the pathology guidelines is inadequate for Raman spectroscopy. This was also seen in IR spectroscopy (Figures 4.4 and 4.5) and is consistent with literature reports [34,55]. Raman spectra obtained from fresh tissue biopsies and cryosections were not contaminated due to the minimal chemical treatments outlined in Chapter 2, Section 2.2, and hence, provided the most reliable diagnostic information.

In order to obtain diagnostic spectra from breast tissues, macrospectroscopic examination using FT-Raman spectroscopy required fresh tissues to be analysed for no longer than 15 min of laser irradiation. Although the S/N ratio of the spectra for longer laser exposure times were better, physical alteration of tissue due to dehydration was observed and, hence, could confound reliable diagnoses, which is unacceptable. Therefore, the research shifted to microspectroscopic approaches where tissues were sectioned to various thicknesses and were placed on different mounting substrates that were also used in histopathology and infrared spectroscopy. In this current study, both CaF_2 windows and mirrors and sample thicknesses of 20 μm produced the best spectral information compared to the standard 5- to 10-μm-thick sections used in histopathology and IR spectroscopy. However, sections on both mounting substrates were not stained for archiving as CaF_2 window is expensive and samples had to be discarded after each spectral analysis. The relatively inexpensive transflective IR slides were used in the later part of this research. However, no signals were recorded on tissue sections (thickness of 5, 10 and 20 μm) mounted on transflective slides using 514-nm laser excitation in this research but these slides were reported to provide quality spectra under these conditions [33].

Differentiation of breast lesions was based on changes in the 1800–800 cm^{-1} spectral region, which contains diagnostic bands similar to those observed in IR spectroscopy. The major spectral differences associated with the malignant breast lesion using 1064-nm laser excitation (Figure 6.9) were an increase in the band intensities attributable to protein (1650, 1300, 936 and 856 cm^{-1}), and DNA (1080 cm^{-1}) and the decrease in intensity of the phenylalanine band (1004 cm^{-1}). Similarly, the amide I and III bands in the malignant IDC cryosections had stronger relative intensities compared to those observed in benign fibrocystic change cryosections when using 514-nm laser excitation (Figure 6.10). The amide I band shifted to lower wavenumber in the spectra of both benign fibrocystic changes cryosections (Figures 6.10B and D) and the low-lipid-malignant cryosection (Figure 6.10C), which was indicative of changes in protein secondary structures from α-helix to β-sheet. However, due to the limited availability of the 830-nm laser during this research, a comparable statistical analysis was not performed.

Raman images of β-carotene were collected to establish the optimal laser defocus percentage for subsequent tissue analysis. The images showed that optimal conditions were achieved with the laser beam being defocused between 5 and 20%. However, photo-degradation was unexpectedly observed in the Raman images obtained with 30 to 50% defocussing. The laser beam is expanded with larger laser defocus percentages to reduce the likelihood of photo-degradation as there is less laser power per unit area on the sample compared to 0% defocused laser beam.

Lobular carcinoma *in-situ* is a rare breast condition (2 samples). It is usually occult in the mammogram and has been reported to be found in association with other invasive carcinomas [49]. With such low detection rates, there is only limited research on this pre-invasive carcinoma using single-point spectroscopy [11]. Therefore, a better understanding of this condition is important for histopathology and vibrational spectroscopy. Raman imaging was employed to characterise the presence of tumour cells within a lobule of a LCIS cryosection using 514-nm laser excitation with 20% defocused laser beam (lobule dimension: ~128 × 128 μm^2, Raman image dimension: ~58 × 58 μm^2). The cellular components were imaged based on amide I band, lipid $\delta(CH_2)$ band and the ring-breathing mode of carotenoid (Figure 6.12). The nuclei had the strongest intensity of these components while the cytoplasm/intralobular stroma had relatively less intensity. As a result, Raman imaging could be used to investigate small lesions at a cellular level.

Similarly, micro-calcification is an important morphological feature for histopathology in order to determine the nature of the breast lesion. There are two types of micro-calcifications found in breast diseases: calcium oxalate dihydrate and calcium hydroxyapatite. The latter is usually associated with malignancy [2,53]. Although Raman imaging had been used to identify calcification in atherosclerotic plaque and brain tumour, there was no imaging reported on breast calcification [53,56]. A suspicious micro-calcification was located within the duct of an IDC breast cryosection. Two significant bands, at 1044 and 961 cm^{-1}, are assigned to calcium hydroxyapatite [1,2,11]. The Raman images constructed using these two bands displayed relatively high intensities of these deposited biominerals in the centre of the area of micro-calcification. The result demonstrated that Raman imaging could be used to identify the formation of micro-calcification and proper medical treatment could be sought.

6.6 REFERENCES

[1] K. Shafer-Peltier, A. Haka, J. Motz, M. Fitzmaurice, R. Dasari, M. Feld, *J. Cell. Biochem.* 39 (2002) 125-137.

[2] K. Shafer-Peltier, A. Haka, M. Fitzmaurice, J. Crowe, J. Myles, R. Dasari, M. Feld, *J. Raman Spectrosc.* 33 (2002) 552-563.

[3] D. Naumann, *Appl. Spectrosc. Rev.* 36 (2001) 239-298.

[4] H. Morjani, A. Beljebbar, G. D. Sockalingum, J. F. Angiboust, M. Polissiou, M. Manfait, *Spectrosc. Biol. Mol.: Mod. Trends, [Eur. Conf.], 7th* (1997) 399-400.

[5] A. Mizuno, H. Kitajima, K. Kawauchi, S. Muraishi, Y. Ozaki, *J. Raman Spectrosc.* 25 (1994) 25-29.

[6] Y. Min, T. Yamamoto, E. Kohda, T. Ito, H. Hamaguchi, *J. Raman Spectrosc.* 36 (2005) 73-76.

[7] R. Manoharan, Y. Wang, M. S. Feld, *Spectrochim. Acta, Part A* 52A (1996) 215-249.

[8] R. Manoharan, Y. Wang, N. Boustany, J. F. Brennan, J. J. Baraga, R. R. Dasari, J. Van Dam, S. Singer, M. S. Feld, *Proc. SPIE-Int. Soc. Opt. Eng.* 2328 (1994) 128-132.

[9] J. Kneipp, T. C. Bakker Schut, M. Kliffen, M. Menke-Pluijmers, G. J. Puppels, *Vib. Spectrosc.* 32 (2003) 67-74.

[10] J. Kneipp, *Vibrational Spectra from Tissues, Cells and Sub-Cellular Structures*. In 6th Australian Conference on Vibrational Spectroscopy, ACOVS6: The University of Sydney, 2005.

[11] A. Haka, K. Shafer-Peltier, M. Fitzmaurice, J. Crowe, R. Dasari, M. Feld, *Cancer Res.* 62 (2002) 5375-5380.

[12] C. J. Frank, D. C. B. Redd, T. S. Gansler, R. L. McCreery, *Anal. Chem.* 66 (1994) 319-326.

[13] C. J. Frank. *Applications of near-Infrared Raman Spectroscopy to the Study of Human Tissue Specimens (Breast Cancer)*, 1994.

[14] M. S. Feld, R. Manoharan, J. Salenius, J. Orenstein-Carndona, T. J. Romer, J. F. Brennan, III, R. Dasari, Y. Wang, *Proc. SPIE-Int. Soc. Opt. Eng.* 2388 (1995) 99-104.

[15] D. C. B. Redd, Z. C. Feng, K. T. Yue, T. S. Gansler, *Appl. Spectrosc.* 47 (1993) 787-791.

[16] T. R. Hata, T. A. Scholz, I. V. Ermakov, R. W. McClane, F. Khachik, W. Gellermann, L. K. Pershing, *J. Invest. Dermatol.* 115 (2000) 441-448.

[17] C. J. Frank, R. L. McCreery, D. C. B. Redd, *Anal. Chem.* 67 (1995) 777-783.

[18] Y. Yazdi, N. Ramanujam, R. Lotan, M. F. Mitchell, W. Hittelman, R. Richards-Kortum, *Appl. Spectrosc.* 53 (1999) 82-85.

[19] J. Tong, M. Meurens, H. Noel, *Near Infrared Spectrosc.: Future Waves, Proc. Int. Conf. Near Infrared Spectrosc., 7th* (1996) 334-336.

[20] K. K. M. Tam, R. S. Armstrong, E. A. Carter, C. E. Mountford, U. Himmelreich, *Raman Spectroscopic Study of Lipid Content in Diseased Human Breast Tissues*, 4th ACOVS, Queensland University of Technology, 2001, 106-107.

[21] K. Shafer-Peltier, R. Manoharan, R. Dasari, M. S. Feld, G. R. Harrison, J. Myles, J. Crowe, M. Fitzmaurice, *Book of Abstracts, 216th ACS National Meeting, Boston, August 23-27* (1998) HYS-344.

[22] B. Schrader, S. Keller, T. Loechte, S. Fendel, D. S. Moore, A. Simon, J. Sawatzki, *J. Mol. Struct.* 348 (1995) 293-296.

[23] B. Schrader, B. Dippel, S. Fendel, S. Keller, T. Lochte, M. Riedl, R. Schulte, E. Tatsch, *J. Mol. Struct.* 408–409 (1997) 23-31.

[24] M. D. Schaeberle, V. F. Kalasinsky, J. L. Luke, E. N. Lewis, I. W. Levin, P. J. Treado, *Anal. Chem.* 68 (1996) 1829-1833.

[25] R. P. Rava, J. J. Baraga, M. S. Feld *Systems and Methods of Molecular Spectroscopy to Provide for the Diagnosis of Tissue.* US Patent: 9215008, 1992.

[26] J. D. Pasteris, B. Wopenka, J. J. Freeman, V. L. Young, H. J. Brandon, *Am. Mineral.* 84 (1999) 997-1008.

[27] L. H. Kidder, V. F. Kalasinsky, J. L. Luke, I. W. Levin, E. N. Lewis, *Nat. Med. (N. Y.)* 3 (1997) 235-237.

[28] C. J. Frank, *Proc. SPIE-Int. Soc. Opt. Eng.* 3608 (1999) 30-36.

[29] M. S. Feld, *Spectrosc. Biol. Mol.: New Dir., Eur. Conf., 8th* (1999) 531-533.

[30] R. N. Baker, J. Smith, K. D. Rogers, N. Stone, *Proc. SPIE-Int. Soc. Opt. Eng.* 6093 (2006) 60911-60931.

[31] R. Alfano, C. Liu, W. Sha, H. Zhu, D. Akins, J. Cleary, R. Prudente, E. Clemer, *Lasers Life Sci.* 4 (1991) 23-28.

[32] J. Radford, *Histopathology Techniques.* 7th ed., Department of Histopathology, The University of Sydney, Sydney, 2002.

[33] E. Faolain, M. Hunter, J. Byrne, P. Kelehan, M. McNamara, H. Byrne, F. Lyng, *Vib. Spectrosc.* 38 (2005) 121-127.

[34] E. Faolain, M. Hunter, H. Byrne, P. Kelehan, H. Lambkin, J. Byrne, F. Lyng, *J. Histochem. Cyctochem.* 53 (2005) 121-129.

[35] G. Socrates, *Infrared and Raman Characteristic Group Frequencies: Tables and Charts.* 3rd ed., John Wiley & Sons Ltd., Chichester, 2001.

[36] K. K. M. Tam. *Spectroscopic Diagnosis of Human Female Breast Diseases,* B.Sc. Hons. Thesis, The University of Sydney, 2001.

[37] T. Kitagawa, S. Hirota, *Raman Spectroscopy of Proteins in Handbook of Vibrational Spectroscopy.* 1st ed., John Wiley & Sons Ltd., Chichester, 2002, Vol. 5, 3426-3446.

[38] M. G. Shim, L. Wong, N. Marcon, B. Wilson, *Photochem. Photobiol.* 63 (1996) 662-671.

[39] B. Schrader, *Infrared and Raman Spectroscopy: Methods and Applications.* 1st ed., VCH Publishers, New York, 1995.

[40] M. Jackson, H. H. Mantsch, *Pathology by Infrared and Raman Spectroscopy in Handbook of Vibrational Spectroscopy.* 1st ed., John Wiley & Sons, Chichester, 2002, Vol. 5, 3227-3245.

[41] H. Fabian, W. Mäntele, *Infrared Spectroscopy of Proteins in Handbook of Vibrational Spectroscopy.* 1st ed., John Wiley & Sons, Chichester, 2002, Vol. 5, 3399-3425.

[42] R. Dukor, *Vibrational Spectroscopy in the Detection of Cancer in Handbook of Vibrational Spectroscopy.* 1st ed., John Wiley & Sons, Chichester, 2002, Vol. 5, 3335-3361.

[43] Z. Huang, A. McWilliams, H. Lui, D. McLean, S. Lam, H. Zeng, *Int. J. Cancer* 107 (2003) 1047-1052.

[44] Y. Guan, E. N. Lewis, I. Levin, *Biomedical Applications of Raman Spectroscopy: Tissue Differentiation and Potential Clinical Usage in Analytical Applications of Raman Spectroscopy.* Blackwell Science Ltd., Oxford, 1999, 276-327.

[45] E. Hanlon, R. Manoharan, T. Koo, K. Shafer, J. Motz, M. Fitzmaurice, J. Kramer, I. Itzkan, R. Dasari, M. Feld, *Phys. Med. Biol.* (2000) R1-R59.

[46] G. McKee, *Cytopathology of the Breast.* 1st ed., Oxford University Press, Boston, 2002.

[47] Thermo-Informatics, *Software: Grams/32 Version 4.14 Level Ii,* 1999.

[48] *Oxford Concise Medical Dictionary.* 6th ed., Oxford University Press, Oxford, 2003.

[49] E. Frykberg, *Breast J.* 5 (1999) 287-350.

[50] S. Silverberg, *Atlas of Breast Pathology.* Elsevier Science (USA), Philadelphia, 2002.

[51] The Susan G. Komen Breast Cancer Foundation (USA), http://www.komen.org 2005.

[52] M. Radi, *Arch. Pathol. Lab. Med.* 113 (1989) 1367.

[53] S. Koljenovic, L. Choo-Smith, T. Schut, J. Kros, H. van den Berge, G. J. Puppels, *Lab. Invest.* 82 (2002) 1265-1277.

[54] J. F. Brennan, III, Y. Wang, R. R. Dasari, M. S. Feld, *Appl. Spectrosc.* 51 (1997) 201-208.

[55] E. O. Faolain, M. B. Hunter, J. M. Byrne, P. Kelehan, H. J. Byrne, F. M. Lyng, *Proc. SPIE-Int. Soc. Opt. Eng.* 5826 (2005) 25-36.

[56] W. van de Poll Sweder, J. Delsing, J. Jukema, H. Princen, L. Havekes, G. J. Puppels, A. van der Laarse, *Atherosclerosis* 164 (2002) 65-71.

7. CONCLUSIONS AND FURTHER EXPERIMENTS

7.1 CONCLUSIONS

Breast cancer affects one in eleven women by the age of 75 in Australia [1]. Triple assessment, a combination of medical examination, X-ray mammography, and a fine-needle aspirated biopsy, is the current gold standard for diagnosis [2]. However, X-ray mammography is not recommended for women under the age of 50 due to the constant changes of the breast content during menstrual cycles. This often results in false positives and negatives that may require the patient to undergo unnecessary further analyses [2]. Therefore, alternative techniques that improve the diagnostic outcomes are necessary. Previously, ^{1}H MRS distinguished benign lesions and malignant lesions with a sensitivity and specificity of 95 and 96 per cent, respectively [3]. However, the MRS technique is destructive and samples cannot be classified by histopathology after spectroscopic analysis [3].

The aim of this research was to develop a common methodology for breast cancer diagnosis that was non-destructive and allowed samples to be analysed in parallel using vibrational spectroscopy and the gold standard of histopathology. One hundred breast lesions (43 benign, 7 *in-situ* carcinomas, 47 malignant, and 3 histopathologically unclassified) were obtained from routine needle or surgical biopsies with appropriate ethics approvals. In order to conduct a parallel study, these samples were prepared according to the standard histopathological procedures, and were subsequently investigated using vibrational spectroscopy on the basis of: optimisation, characterisation and differentiation.

7.1.1 TISSUE OPTIMISATION FOR VIBRATIONAL SPECTROSCOPY

In order to develop a robust and reliable protocol for routine cancer diagnosis, it was essential to evaluate the histopathological protocols with vibrational spectroscopy. Specific areas of interest included: tissue fixation, processing, and sectioning, together with the substrates for mounting.

In this research, both fresh and formalin-fixed paraffin-embedded (FFPE) biopsies were collected for FTIR and Raman microspectroscopic analyses. The FTIR spectra demonstrated that cryosectioned fresh samples yielded the best diagnostic results compared to those obtained when the samples were prepared as FFPE sections. This was on the proviso that the cryosections were analysed within a week after sectioning, otherwise the tissues were degraded as a result of autolysis and putrefaction [4]. On the other hand, FFPE sections were fixed allowing for long-term storage and retrospective analysis could also be undertaken [4]. No significant spectral differences were observed between the spectra collected from the same FFPE sections three years apart.

Samples were mounted on a variety of substrates including, CaF_2 windows, transflective slides and metal-based mirrors. The most appropriate substrate for Raman spectroscopy was a CaF_2 window, which is also the preferred substrate for FTIR spectroscopy. However, the cost of such windows is prohibitive and the transflective slides were found to be the best high-throughput alternative.

An additional parameter that needed to be considered for Raman spectroscopy was the selection of the most appropriate laser excitation wavelength (514, 647, 752, 830 or 1064 nm). The strong fluorescence background generated using 647- and 752-nm excitation masked the Raman features of biological components. Subsequent research focussed on the use of 514-, 830-, and 1064-nm excitation wavelengths, which produced different spectral information.

7.1.2 DIFFERENTIATION OF BREAST LESIONS

Spectroscopic differentiation of breast lesions that cannot be discriminated by histopathology is important for improved cancer diagnosis. Pathologists diagnose malignancy by examining a sample under a light microscope to look for morphological changes that have occurred such as cell proliferation and ductal enlargement. Long before these changes eventuate, the cells have undergone numerous biochemical changes in epithelial components within and around the ducts or lobules. These biochemical changes occur at a molecular level and result in increased intensities or changes in the lineshapes or positions of diagnostic IR and Raman bands (proteins, nucleic acids, and lipids).

7.1.2.1 INFRARED SPECTROSCOPY

The present study demonstrates that benign, *in-situ* carcinomas and malignant lesions in breast tumour tissue sections can be differentiated based on their IR spectra. Three discrete spectral regions were analysed, of which the fingerprint region (1700-1000 cm^{-1}) provided the diagnostic bands that were commonly used for tissue differentiation in this research and that published in the literature [2-28]. This is the first IR spectroscopic study undertaken that specifically focuses on specific types of lesions using FPA-FTIR imaging and SR-FTIR mapping (i.e., fibrocystic change, DCIS and IDC).

The visual examination of infrared spectra from breast lesions was proved to be unreliable, due to similarities among spectra from different types of lesions, as shown in Section 5.2. Curve-fitting and second-derivatives were applied to the raw IR data to resolve the bands attributed to biomolecules (proteins, lipids, and nucleic acids). Principal component analysis and hierarchical cluster analysis successfully separated spectra obtained from malignant lesions and those obtained from benign lesions using single-point microspectroscopy over the spectral region of 1700-1000 cm^{-1}. A PCA scores plot clearly distinguished the spectra of fibrocystic changes from the spectra of benign fibroadenomas and intraduct papillomas. Although the spectra of IDC and ILC sections were not clearly separated using the raw data, a PCA scores plot using second-derivative spectra distinguished both IDC and ILC lesions from each other. The use of PCA for classification of benign and malignant sub groups has not been reported previously. Similarly, HCA was able to differentiate benign and malignant breast lesions using the combination of spectral regions. Two breast lesions without prior knowledge to the histopathological outcomes were classified as fibrocystic changes using HCA, which were subsequently correlated with the histopathological results. The IR spectra of two pre-classified malignant IDC tissue sections were clustered using HCA within the DCIS (D) sub-group. These two lesions were subsequently stained after IR analysis and histopathology results confirmed that these lesions were, in fact, high-grade DCIS in which cell proliferation was in the process of invading through the basement membrane, as evident from the adjacent H&E stained section. The HCA classification and subsequent histopathology examination overruled the original classification of being malignant IDC. Spectral classification of benign, *in-situ* carcinomas and malignant lesions using PCA and HCA were achieved with overall sensitivities of 100/100 (PCA/HCA), 100/78 and 92/91%, and specificities of 100/94, 96/96 and 100/100%, respectively. The results were comparable with previous research using ^{1}H

MRS, which distinguished malignant tissues from benign lesions with sensitivity and specificity of 95 and 96%, respectively [3].

7.1.2.2 RAMAN SPECTROSCOPY

Differentiation of breast lesions with Raman spectroscopy was based on changes in the 1800–800 cm^{-1} spectral region, which contains diagnostic bands similar to those observed in IR spectroscopy. The major spectral differences associated with the malignant breast lesion using 1064-nm laser excitation (Figure 6.9) were an increase in the band intensities attributable to protein (1650, 1300, 936 and 856 cm^{-1}), and DNA (1080 cm^{-1}) and the decrease in intensity of the phenylalanine band (1004 cm^{-1}). Similarly, the amide I and III bands in the spectra from the malignant IDC cryosections had stronger relative intensities compared to those observed in benign fibrocystic change cryosections when using 514-nm laser excitation (Figure 6.10). The amide I band shifted to lower wavenumber in the spectra of both benign fibrocystic changes cryosections and the low-lipid-malignant cryosection, which was indicative of changes in protein secondary structures from α-helix to β-sheet. However, the sensitivity and specificity of various breast diseases was not determined due to the insufficient samples examined using Raman spectroscopy,

7.1.3 CHARACTERISATION OF BREAST LESIONS

Characterisation of specific breast diseases, such as ductal carcinoma *in-situ* (DCIS) and lobular carcinoma *in-situ* (LCIS), are important to medical diagnosis since both conditions are pre-cursors of invasive breast cancer [5,6].

7.1.3.1 FPA-FTIR IMAGING

Focal-plane array infrared imaging is a powerful technique, which was used for characterisation of tissue sections. In this present study, non-hierarchical cluster analysis (k-means clustering and fuzzy c-means clustering) and principal component analysis were applied to the imaging data from a lesion-free breast duct, a low-grade cribriform DCIS and an invasive ductal carcinoma. When referenced to the H&E stained sections, these statistical analyses differentiated cellular components (e.g., tumour cells, adipose and connective tissues) by grouping similar spectra together, and assigning a specific colour to a cluster for each tissue component.

Prior to choosing the number of clusters used in multivariate statistical analysis, the tissue components had to be identified from the adjacent H&E stained section together with the histopathology results. The use of insufficient or an excess number of clusters will not account for the true spectral variation, and thus the images cannot be compared reliably with histopathology results. The present study demonstrated that five clusters were sufficient to delineate the major tissue components within the breast lesions (normal and diseased) and the k-means clustering technique yielded the fastest computational time and simplest interpretation compared to those obtained using fuzzy c-means clustering and principal component analysis.

7.1.3.2 SR-FTIR MAPPING

Synchrotron-radiation-based FTIR spectroscopy is a relatively new technique, which can be applied for tissue diagnosis. There is limited research on breast tissue characterisation with this technique due to accessibility problems and relatively long mapping times [7,8]. For example, a low-grade cribriform DCIS section (350 x 350 μm^2) was mapped and imaged with SR-FTIR and FPA-FTIR microspectroscopy, respectively. The collection time required for the same area under analysis was ten hours using SR-FTIR microspectroscopy due to a lack of an array detector on the beamline used compared with less than five minutes using FPA-FTIR imaging.

In the present study, a benign fibrocystic change, a DCIS and an IDC deparaffinised FFPE section were mapped. The resulting functional group maps displayed the distribution of proteins, lipids and nucleic acids on the tissue sections, and were highly correlated to the areas of particular interest in each lesion (i.e., tumour cells, ducts and lobules) of the H&E stained sections.

7.1.3.3 RAMAN IMAGING

Lobular carcinoma *in-situ* is a rare breast condition (2 samples). It is usually occult in the mammogram and has been reported to be found in association with other invasive carcinomas [6]. With such low detection rates, there is only limited research on this pre-invasive carcinoma using single-point spectroscopy [9]. Therefore, better morphological and spectroscopic knowledge of LCIS is beneficial for diagnostics based on both histopathology and vibrational spectroscopy. Raman imaging was employed to characterise the presence of tumour cells within a lobule of a LCIS cryosection using 514-nm laser excitation with 20% defocused laser beam (lobule dimension: ~128 × 128 μm^2, Raman image dimension: ~58 × 58 μm^2). The cellular components were imaged based on the amide I band, lipid $\delta(CH_2)$ band and the band due to the ring-breathing mode of carotenoids (Figure 6.12). Cell nuclei gave the strongest Raman band intensities for these components while the cytoplasm/intralobular stroma resulted in a lower relative intensity of these bands. As a result, Raman imaging could be used to investigate small lesions at a cellular level.

The presence of a micro-calcification is another indicator of a lesion being cancerous, and it is also occult on a mammogram [10,11]. A Type II micro-calcification (calcium hydroxyapatite) was identified in the Raman spectrum from an IDC lesion, which was consistent with reports in the literature of its association with IDC [9]. Raman images based on the bands due to the $\nu(PO_4^{2-})$ modes, at 1044 and 961 cm^{-1}, illustrated that the calcification was denser in the centre of the lesion and gradually decreased in concentration towards the outer region. It is proposed that imaging could be used to establish the seeding process and progress of micro-calcification.

7.2 FURTHER EXPERIMENTS

7.2.1 COMBINED PATHOLOGY, MRS AND VIBRATIONAL SPECTROSCOPY

One of the challenges in this study was to improve the workflow of all experimental work in pathology, MRS and vibrational spectroscopy, i.e., to decrease the diagnostic time, to improve its reliability, and/or to reduce ultimately the waiting time for the patients. The current strategy is time-consuming and revolves around a complicated network of experiments (Figure 7.1). Therefore, a tighter integration and targeted experimental workflow is sought to reduce the experimental time for acceptance of using vibrational spectroscopy in medical diagnosis.

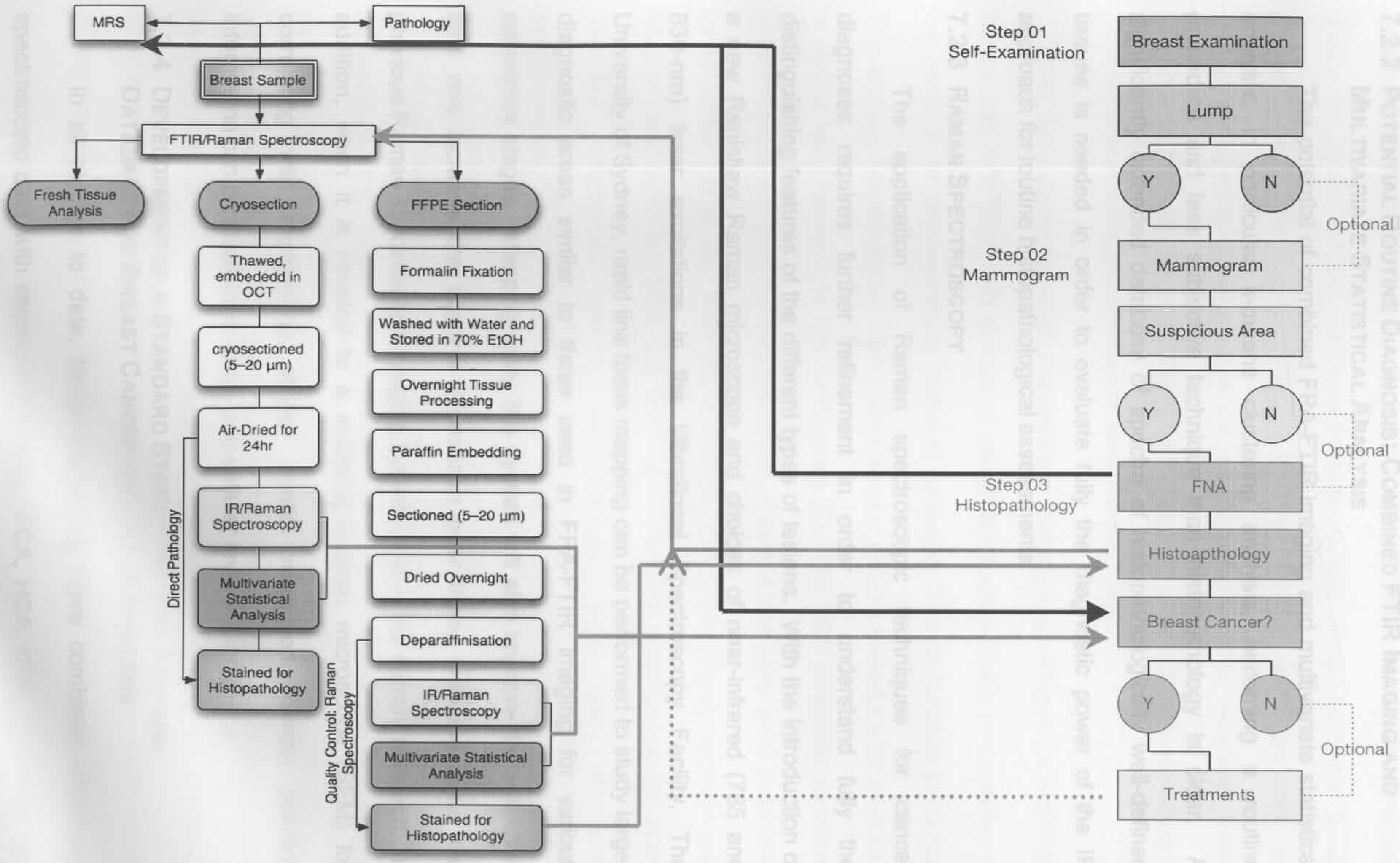


Figure 7.1. Multi-disciplinary strategy for incorporating MRS, and IR and Raman spectroscopies into routine breast cancer diagnosis.

7.2.2 POTENTIAL ROUTINE DIAGNOSIS: COMBINED FTIR IMAGING AND MULTIVARIATE STATISTICAL ANALYSIS

The potential of combined FPA-FTIR imaging and multivariate statistical analysis, in particular k-means clustering analysis, becoming a routine procedure and less subjective technique than histopathology is clear. A significantly extended database of spectra of histopathologically well-defined tissues is needed in order to evaluate fully the diagnostic power of the IR approach for routine histopathological assessments.

7.2.3 RAMAN SPECTROSCOPY

The application of Raman spectroscopic techniques for cancer diagnoses requires further refinement in order to understand fully the distinguishing features of the different types of lesions. With the introduction of a new Renishaw Raman microscope and choices of near-infrared (785 and 830-nm) laser excitations in the Vibrational Spectroscopy Facility, The University of Sydney, rapid line focus mapping can be performed to study larger diagnostic areas similar to those used in FPA-FTIR imaging for various cancerous stages. Mapping using this system will also improve the collection time and increase the spectral information over those obtained using the previous Raman spectroscopic imaging system used in the current research. In addition, when it is coupled to a scanning-electron microscope (SEM) for correlating with morphology, higher spatial resolution Raman spectral information can be obtained to study sub-cellular level of the tissues.

7.2.4 DEVELOPMENT OF A STANDARD STATISTICAL ANALYSIS AND DATABASE FOR BREAST CANCER AND OTHER CANCERS

In all studies to date, biospectroscopists have combined vibrational spectroscopic data with chemometrics (e.g., PCA, HCA, KMC, and FCM) for tissue characterisation and discrimination. These statistical analyses distinguish subtle spectral variations between cellular components and group similar spectra with the same tissue component (e.g., epithelial cells) by assigning a specific colour pixel on the resulting cluster map. However, these statistical techniques are usually difficult to carry out and interpret, especially if inappropriate clusters were chosen to describe the image.

A three-stage statistical classification strategy was designed specifically for analysis of MRS and IR spectra of biological samples (fluids and tissues) at the Institute for Biodiagnostics, National Research Council Canada [3,12,13]. The strategy has developed highly accurate and reliable classifiers for a variety of clinical MRS research by the IMRR group [3]. The SCS schematic workflow is summarised in Figure 7.2 and the details of each stage have been published by Mountford, *et al.* [3]. With the advantage of providing highly accurate and reliable classifiers, the SCS can become a standard classification for inter-discipline diagnostics among histopathology, MRS and vibrational spectroscopy.

In conclusion, infrared or Raman microspectroscopy using the specialised techniques of mapping and imaging together with SCS could demonstrate the potential capabilities of the non-subjective diagnostic outcomes to medical practitioners. Notwithstanding, the strategy still requires refinement, especially in the area of FPA-FTIR imaging, SR-FTIR mapping and near-infrared Raman microspectroscopy. Therefore, a strong collaboration, among medical experts (surgeons and pathologists) and spectroscopists (MRS, IR and Raman), is essential to develop methodologies for routine diagnosis with high sensitivity and selectivity.

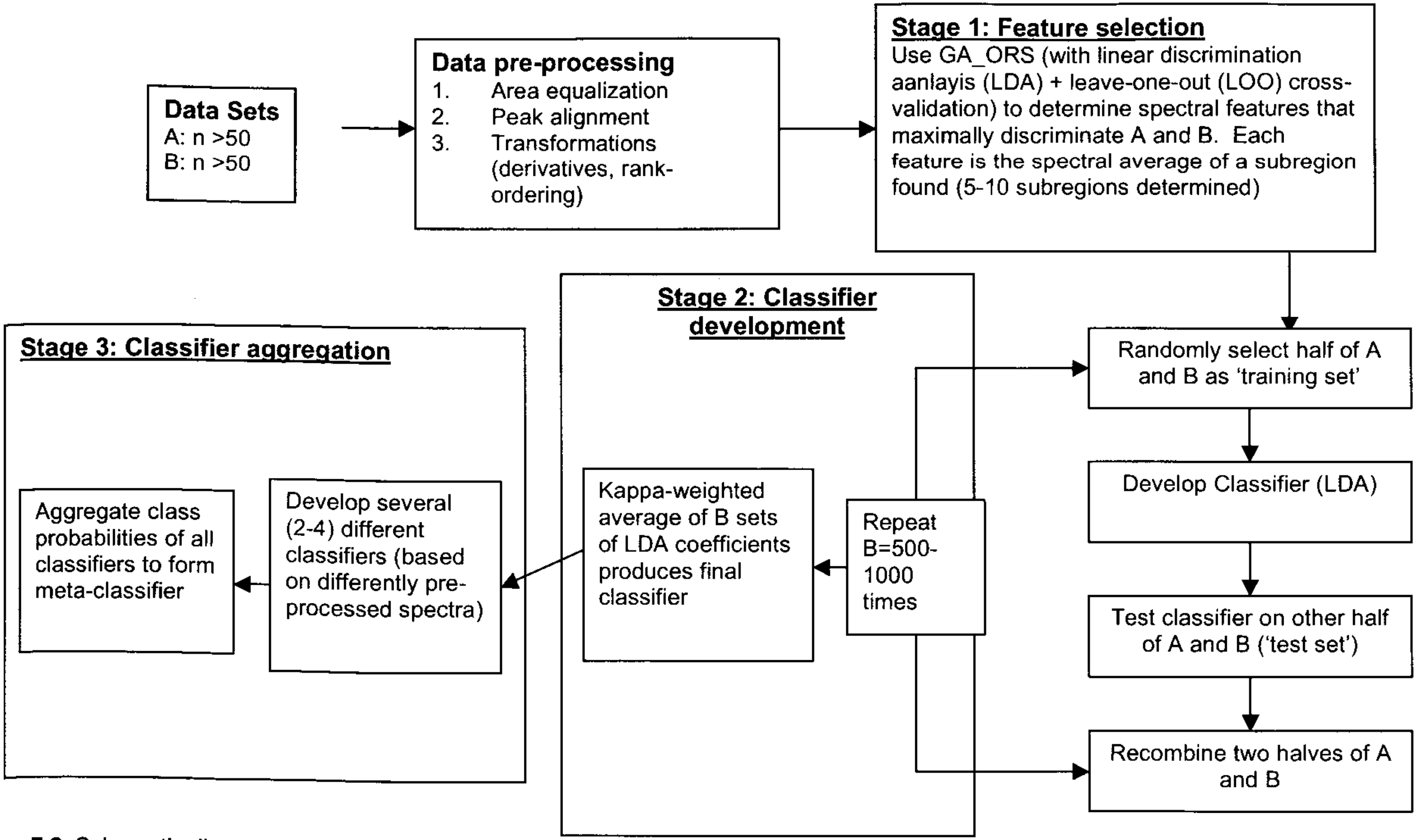


Figure 7.2. Schematic diagram of statistical cluster strategy currently used by Mountford *et al.* [3] for magnetic resonance research.

7.3 REFERENCES

[1] C. Stevenson, *Breastscreen Australia Monitoring Report 2001-2002*. Australian Institute for Health and Welfare, Canberra, 2005.

[2] G. McKee, *Cytopathology of the Breast*. 1st ed., Oxford University Press, Boston, 2002.

[3] C. Lean, R. Somorjai, I. Smith, P. Russell, C. Mountford, *Ann. Rev. NMR. S.* 48 (2002) 71-111.

[4] J. Radford, *Histopathology Techniques*. 7th ed., The University of Sydney, Sydney, 2002.

[5] H. Burstein, K. Polyak, J. Wong, S. Lester, C. Kaelin, *N. Engl. J. Med.* 350 (2004) 1430-1441.

[6] E. Frykberg, *Breast J.* 5 (1999) 287-350.

[7] C. Liu, Y. Zhang, X. Yan, X. Zhang, C. Li, W. Yang, D. Shi, *J. Lumin.* 119-120 (2006) 132-136.

[8] R. N. Baker, J. Smith, K. D. Rogers, N. Stone, *Proc. SPIE-Int. Soc. Opt. Eng.* 6093 (2006) 60930X/60931-60930X/60911.

[9] A. Haka, K. Shafer-Peltier, M. Fitzmaurice, J. Crowe, R. Dasari, M. Feld, *Cancer Res.* 62 (2002) 5375-5380.

[10] R. Blamey, A. Evans, I. Ellis, R. Wilson, *Atlas of Breast Cancer*. 1st ed., Merit Publishing International, Basingstoke, 1994.

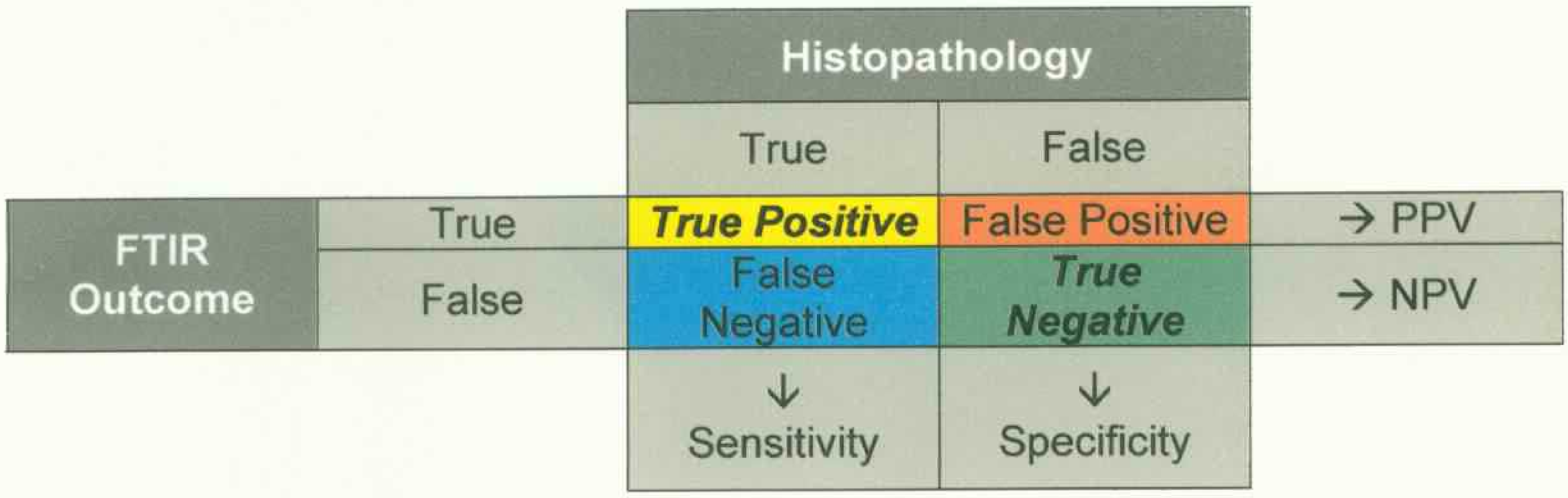
[11] N. Weidner, *Breast in Modern Surgical Pathology.* Saunders, 2003, 540-571.

[12] M. Jackson, J. R. Mansfield, B. Dolenko, R. L. Somorjai, H. H. Mantsch, *Cancer Detect. Prev.* 23 (1999) 245-253.

[13] C. Mountford, P. Russell, P. Malycha, L. Gluch, R. Somorjai, C. Lean, B. Barrraclough, D. Gillett, U. Himmelreich, B. Dolenko, E. Nikulin, C. Smith, *Br. J. Surgery* 88 (2001) 1234-1240.

APPENDIX A – FORMULAE FOR SENSITIVITY AND SPECIFICITY

Figure A.1. Determination of the sensitivity and specificity from FTIR diagnostics and histopathology. Arrows represent the parameters used to calculate the specific diagnostic outcomes (sensitivity, specificity, PPV, and NPV). PPV: positive predictive value and NPV: negative predictive value used in Chapter 5, Section 5.2.4.



		Histopathology		
		True	False	
FTIR Outcome	True	***True Positive***	False Positive	→ PPV
	False	False Negative	***True Negative***	→ NPV
		↓ Sensitivity	↓ Specificity	

Sensitivity

$$\text{Sensitivity} = \frac{\text{Number of True Positives}}{\text{Number of True Positives} + \text{Number of False Negatives}}$$

Specificity

$$\text{Specificity} = \frac{\text{Number of True Negatives}}{\text{Number of True Negatives} + \text{Number of False Positives}}$$

Positive predictive value (PPV)

$$\text{PPV} = \frac{\text{Number of True Positives}}{\text{Number of True Positives} + \text{Number of False Positives}}$$

Negative predictive value (NPV)

$$\text{NPV} = \frac{\text{Number of True Negatives}}{\text{Number of True Negatives} + \text{Number of False Negatives}}$$